THE PHYSICS AND CHEMISTRY OF SOLIDS

AN INTERNATIONAL JOURNAL

VOLUME 1 1956/1957

Editor-in-Chief **HARVEY BROOKS** *Cambridge, Mass.*

Associate Editors
H. B. G. CASIMIR *Eindhoven*
G. J. DIENES *Brookhaven*
JACQUES FRIEDEL *Paris*
E. M. LIFSHITZ *Moscow*
L. D. LANDAU *Moscow*

Editorial Advisory Board

U.S.A.: L. Apker, Schenectady · J. Bardeen, Illinois · D. S. Billington, Oak Ridge · Leo Brewer, Berkeley, Calif. · N. Cabrera, Virginia · I. Estermann, Washington, D.C. · J. C. Fisher, Schenectady · Conyers Herring, New Jersey · A. R. von Hippel, Cambridge, Mass. · J. H. Hollomon, Schenectady · W. V. Houston, Houston · H. B. Huntington, Troy · John G. Kirkwood, Yale · K. Lark-Horowitz, Purdue · A. W. Lawson, Chicago · B. T. Matthias, New Jersey · Joseph E. Mayer, Chicago · David Pines, Princeton R. Smoluchowski, Pittsburgh · D. Turnbull, Schenectady · J. H. Van Vleck, Harvard · B. E. Warren, Cambridge, Mass. · Eugene P. Wigner, Princeton · B. H. Zimm, Schenectady.

United Kingdom: M. Born, Edinburgh · A. Charlesby, Cambridge · W. J. Dunning and F. C. Frank, Bristol H. K. Henisch, Reading · H. Jones, London · J. W. Mitchell, Bristol · A. B. Pippard, Cambridge · M. H. L. Pryce, Bristol · F. C. Tompkins, London.

France: E. F. Bertaut, Grenoble · A. Guinier, Paris · Louis Néel and Louis Weil, Grenoble.

Germany: Ulrich Dehlinger, Stuttgart · R. Hilsch and G. Leibfried, Göttingen · E. Mollwo, Erlangen H. Pick, Stuttgart · M. von Laue, Berlin · H. Welker, Erlangen.

The Netherlands: A. Michels and G. W. Rathenau, Amsterdam. *Sweden:* J. O. Linde, Stockholm · Per-Olov Löwdin, Uppsala.

Switzerland: G. Busch and P. Scherrer, Zürich. *U.S.S.R.:* G. S. Zhdanov, Moscow.

PERGAMON PRESS **NEW YORK · LONDON · PARIS**

NOTES FOR CONTRIBUTORS

I. GENERAL

1. Submission of a paper to the *Journal of the Physics and Chemistry of Solids* will be taken to imply that it represents original research not previously published (except in the form of an abstract or preliminary report), that it is not being considered for publication elsewhere, and that if accepted it will not be published elsewhere in the same form, in any language, without the consent of the editor-in-chief. It should deal with original research work in the field of the physics and chemistry of solids.

2. Papers should be submitted to the appropriate regional editor (all English-language papers to be sent to the U.S. editor).

3. Papers will be published as quickly as possible after acceptance, and, subject to space being available, should appear in the issue next following, if this is due for publication not earlier than three months after the acceptance date. Short communications under the heading of "Letters to the Editors" will receive priority for publication, and will be published in the issue following receipt, if accepted not later than the beginning of the month preceding publication.

4. Fifty free reprints of each paper are supplied. Additional copies can be obtained at a reasonable cost if ordered when proofs are returned. A reprint order form will accompany first proofs.

II. SCRIPT REQUIREMENTS

1. Papers submitted should be concise and written in a readily understandable style. Scripts should be typed and double spaced and submitted in duplicate to facilitate refereeing.

It will be appreciated if authors clearly indicate any special characters used. An abstract, not exceeding 200 words, should be provided in the language of the paper. French and German papers should be submitted with English abstract and titles, but if this is not possible the abstract will be translated by the publishers. To conserve space, authors are requested to mark less important parts of the paper, such as details of experimental technique, methods, mathematical derivations, etc., for printing in small type. The technical description of methods should be given in detail only when such methods are new. Authors will receive proofs for correction when their papers are first set; page proofs will be sent only when the amount of alteration makes it advisable.

2. Illustrations should not be included in the typescript of the paper, and legends should be typed on a separate sheet. Line drawings which require redrawing should include all relevant details and clear instructions for the draughtsman. If figures are already well drawn it may be possible to reproduce them direct from the originals, or from good photo-prints if these can be provided. It is not possible to reproduce from prints with weak lines. Illustrations for reproduction should normally be about twice the final size required. Photographs should only be included where they are essential.

Tables and figures should be so constructed as to be intelligible without reference to the text. Every table and column should be provided with an explanatory heading. Units of measure must always be clearly indicated. The same data should not be published in both tables and figures. The following standard symbols should be used on line drawings since they are easily available to the printers: ○, ●, +, ×, □, ■, △, ▲, ◇, ◆, ▽, ▼

4. References are indicated in the text by superior numbers in parentheses, and the full reference should be given in a list at the end of the paper in the following form:

1. HILL R. *J. Iron and Steel Inst.* **158,** 177 (1948).
2. PEARSON C. *The Extrusion of Metals* Chapman and Hall, London (1944).

Abbreviations of journal titles should follow those given for *Physics Abstracts*. It is particularly requested that authors' initials, and appropriate volume and page numbers, should be given in every case.

Footnotes, as distinct from literature references should be indicated by the following symbols—*, †, ‡, ‡‡, commencing anew on each page; they should *not* be included in the numbered reference system.

5. Due to the international character of the journal no rigid rules concerning notation and spelling will be observed, but each paper should be consistent within itself as to symbols and units.

VOL.
1
1956/5

OL.
1
6/57

VOL.
1
1956/5

CONTENTS

NUMBERS 1/2

Foreword 1

GERT EHRLICH: The mechanism of chemisorption on metals 3

G. DRESSELHAUS: Effective mass approximation for excitons 14

A. C. DAMASK: Residual resistivity vs. short-range order in Cu_3Au 23

JOHN B. GIBSON: The effect of short-range order on residual resistivity 27

E. P. WOHLFARTH: The electronic properties of nickel-palladium alloys 35

A. I. SCHINDLER, R. J. SMITH, and E. I. SALKOVITZ: Preliminary electrical-resistivity measurements of the nickel-palladium alloy system 39

A. I. SCHINDLER: The Hall effect in the silver-palladium alloy system 42

J. A. HOFMANN, A. PASKIN, K. J. TAUER, and R. J. WEISS: Analysis of ferromagnetic and antiferromagnetic second-order transitions 45

JAMES C. THOMPSON: Resistance and paramagnetism at the superconducting transition 61

E. BURSTEIN, G. PICUS, B. HENVIS, and R. WALLIS: Absorption spectra of impurities in silicon—I. Group-III acceptors 65

G. PICUS, E. BURSTEIN, and B. HENVIS: Absorption spectra of impurities in silicon—II. Group-V Donors 75

E. O. KANE: Energy band structure in *p*-type germanium and silicon . . . 82

HARRY LETAW, Jr.: Thermal acceptors in germanium 100

L. M. MATARRESE and CHIHIRO KIKUCHI: Paramagnetic resonance absorption of Mn^{++} in single crystals of zincblende 117

Book Review:
"Solid-State Physics." Edited by F. SEITZ and D. TURNBULL.
(Reviewed by N. F. MOTT) 128

Announcement 128

NUMBER 3

A. M. CLOGSTON, H. SUHL, L. R. WALKER, and P. W. ANDERSON: Ferromagnetic resonance line width in insulating materials 129

Y. YAFET, R. W. KEYES, and E. N. ADAMS: Hydrogen atom in a strong magnetic field . 137

ROBERT W. KEYES and R. J. SLADEK: Effect of a magnetic field on donor impurity levels in InSb 143

JOHN W. STEWART: Compression of solidified gases to 20,000 kg/cm² at low temperature 146

H. R. PHILIPP and E. A. TAFT: Photoelectric emission from the valence band of cesium iodide 159

I. M. LIFSHITZ, M. YA AZBEL, and V. I. GERASIMENKO: Paramagnetic resonance and polarization of nuclei in metals 164

J. FRIEDEL: Sur la structure de l'uranium métallique 175

B. T. MATTHIAS, E. A. WOOD, E. CORENZWIT, and V. B. BALA: Super-conductivity and electron concentration 188

W. DALE COMPTON and ROBERT J. MAURER: Self-diffusion and electrical conductivity in silver chloride 191

R. R. HAKE and D. E. MAPOTHER: Effect of pressure on the superconducting transition of lead 199

NUMBER 4

H. SUHL: The theory of ferromagnetic resonance at high signal powers . . . 209

HIROSHI SATO and B. S. CHANDRASEKHAR: Determination of the magnetic anisotropy constant K_2 of cubic ferromagnetic substances 228

RICHARD H. BUBE: Analysis of photoconductivity applied to cadmium-sulfide-type photoconductors 234

EVAN O. KANE: Band structure of indium antimonide 249

F. J. BLATT: Scattering of carriers by ionized impurities in semiconductors . . 262

B. R. COLES and J. C. TAYLOR: Sign reversal of the hall effect in rhodium . . 270

D. O. THOMPSON and D. K. HOLMES: The effect of gamma irradiation on the Young's modulus of copper 275

U. DEHLINGER: Zur Elektronentheorie metallischer und halbleitender Strukturen . 279

Letter to the Editors:

RICHARD PETERSEN: Exciton doublet splitting in ionic crystals 284

VOL.
1
1956/5

AUTHOR INDEX

(B) denotes book reveiw

ADAMS, E. N. 137
ANDERSON, P. W. 129
AZBEL', M. YA. 164

BALA, V. B. 188
BLATT, F. J. 262
BUBE, R. H. 234
BURSTEIN, E. 65, 75

CHANDRASEKHAR, B. S. 228
CLOGSTON, A. M. 129
COLES, B. R. 270
COMPTON, W. D. 191
CORENZWIT, E. 188

DAMASK, A. C. 23
DEHLINGER, U. 279
DRESSELHAUS, G. 14

EHRLICH, G. 3

FRIEDEL, J. 175

GERASIMENKO, V. I. 164
GIBSON, J. B. 27

HAKE, R. R. 199
HENVIS, B. 65, 75
HOFMANN, J. A. 45
HOLMES, D. K. 275

KANE, E. O. 82, 249
KEYES, R. W. 137, 143
KIKUCHI, C. 117

LETAW, H., Jr. 100
LIFSHITZ, I. M. 164

MAPOTHER, D. E. 191
MATARRESE, L. M. 117
MATTHIAS, B. T. 188
MAURER, R. J. 191
MOTT, N. F. 128(B)

PASKIN, A. 45
PETERSEN, R. 284
PHILIPP, H. R. 159
PICUS, G. 65, 75

SALKOVITZ, E. I. 39
SATO, H. 228
SCHINDLER, A. I. 39, 42
SLADEK, R. J. 143
SMITH, R. J. 39
STEWART, J. W. 146
SUHL, H. 129, 209

TAFT, E. A. 159
TAVER, K. I. 45
TAYLOR, J. C. 270
THOMPSON, D. O. 275
THOMPSON, J. C. 61

WALKER, L. R. 129
WALLIS, R. 65
WEISS, R. J. 45
WOHLFARTH, E. P. 35
WOOD, E. A. 188

YAFET, Y. 137

OL.
1
6/57

VOL.
1
1956/5

FOREWORD

THE emergence of solid-state physics as a recognized specialty of physics has taken place over a period of many years. A more recent development, stimulated partly by the growth of industrial interest in the field, has been the growing realization of the common interests between physicists and chemists in the problems of solids. It has been felt by many that the coming of age of solid-state science should be recognized by the publication of a journal devoted exclusively to this field. The present journal represents the effort of a private publisher to meet this need. The favorable response of leaders in the field the world over to the suggestion of such a project has encouraged the publisher to believe that the need is a genuine one, and induced the editor to undertake his duties.

While there is certainly a place for purely national journals of general physics, it has seemed to us that a specialized journal has little justification unless it serves to stimulate the exchange of ideas on the widest international basis. Of course, scientists from different countries read each other's national journals, and there is extensive exchange of ideas on a private basis. Nevertheless, we believe that the clear recognition of a special field of interest in a common medium of publication does provide a natural channel for the exchange of ideas and criticism, on the widest basis, which cannot be achieved in any other way.

The journal will publish both theoretical and experimental papers dealing with original research in all aspects of the physical and chemical properties of solids and condensed systems. The principal fields of interest, however, will center on (1) The electronic structure of solids and those physical and chemical properties which derive more or less directly from electronic structure, and (2) The fundamental statistical mechanics of condensed systems.

Such fields as the quantum theory of solids, ferroelectricity, luminescence, semiconductor physics, magnetism, magnetic resonance, and superconductivity, are well recognized as branches of solid-state science and will be heavily represented in the journal. It is the hope of the editors, however, that the journal will also succeed in developing the lines of communication between physicists and chemists in the conventional recognized areas of solid state and individuals working in such areas as the theory and structure of liquids, the physics of high polymers, and adsorption and catalysis. It is our belief that many concepts recently developed in connection with the simpler solids will prove very fruitful in these other areas, and conversely that the problems of these other areas will stimulate the further evolution and exploitation of solid-state concepts.

The focus of the journal will be fundamental rather than applied. For this reason preference will be given to those contributions having the widest significance for the understanding of the properties of solids in general, and emphasis will thus be on work which has some implications for the theory—though by no means necessarily theoretical work.

It is with considerable pride that we offer this first double number of the journal to our readers. We believe the selection of papers to be a good representation of the type of journal we wish to publish. In two respects, however, this first issue falls short of our ultimate hopes and aims. In the first place, owing largely to the problems of initial organization of an international enterprise, the contributions to this first issue are predominantly American. We hope that future numbers will assume an increasingly international character. In the second place, there is not as large a representation as we should like of papers from borderline fields such as polymers and liquids.

Cambridge, Massachusetts HARVEY BROOKS
31 *August* 1956

VOL.
1
1956/5

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 1–13.

THE MECHANISM OF CHEMISORPTION ON METALS

GERT EHRLICH

General Electric Research Laboratory, Schenectady, New York

(*Received* 3 *March* 1956)

Abstract—The available experimental data on the adsorption of diatomic molecules on a bare metal surface are analysed to establish the kinetic processes leading to chemisorption. It is shown that, with the possible exception of hydrogen, molecules physically adsorbed on the surface form a reservoir from which chemisorption can occur.

Kinetic mechanisms involving such a physically adsorbed precursor are developed, both for reaction on a uniform surface as well as for reaction at surface singularities, such as grain boundaries or lattice steps. Analysis of the rate of adsorption of nitrogen on tungsten reveals that during the initial stages the rate of chemisorption is diffusion controlled, and occurs preferentially at heterogeneities of atomic dimensions, which are tentatively identified with lattice steps on the surface.

1. INTRODUCTION

ALTHOUGH the chemisorption of gases on metals has been under investigation for almost half a century, the experimental difficulties involved have been so great as to limit severely the study of the kinetics of this process. So far, there appear to be only three systems for which the rate of adsorption has been carefully studied under conditions designed to minimize extraneous effects due to impurities and contaminants—the system cesium on tungsten by TAYLOR and LANGMUIR,[1] oxygen on tungsten by MORRISON and ROBERTS,[2] and nitrogen on tungsten by BECKER and HARTMAN.[3]

From the first of these has arisen the classical picture of chemisorption of atoms—impingement of an atom from the vapor upon a surface, followed by diffusion in a loosely held layer over atoms already adsorbed, and ending in chemisorption once a vacant site is reached. Except for a small number of active spots, the polycrystalline surface was found to be uniform, not only in bringing about adsorption of every atom incident upon it, but also in the thermodynamic properties of the adsorbed species. MORRISON and ROBERTS, from their rate measurements, concluded that for molecules also one of the steps leading to chemisorption involved a loosely bound layer of adsorbed entities, which could diffuse over the surface to the site at which chemisorption occurs, and BECKER appears to have made this assumption implicitly in interpreting his data. Partly because even the limited data available have not been satisfactory analysed, we still lack beyond this as much as a qualitative understanding of the phenomena of chemisorption —the steps leading to it, the effects of surface structure upon the rate, and even an understanding of the state of binding referred to by this name. In the following, we shall examine in some detail the available experimental facts, with a view to elucidating the rôle of a precursor adsorbed on the surface. Using this concept, we shall then present several models for the kinetic processes involved in chemisorption, and subject these to comparison with experiment, particularly with the aim of establishing the importance of active surface configurations.

2. THE PRECURSOR TO CHEMISORPTION

It was LANGMUIR[4] who suggested in 1929 that molecules of oxygen adsorbed on top of oxygen adatoms (present on tungsten) could take part in chemisorption, but experimental evidence in support of such a general mechanism was presented by TAYLOR and LANGMUIR[1] only somewhat later. In the study of the adsorption of cesium on tungsten, they found that the sticking coefficient, defined as the fraction of the number of molecules incident upon unit area in unit time that become chemisorbed, remained at unity even at coverages as high as 0·98 of a monolayer, for which a majority

of the atoms colliding with the surface would strike filled sites. From this they concluded that chemisorption on partly filled surfaces did not occur solely through direct collision of a cesium atom from the vapor with an available site on the surface, but rather involved prior condensation of cesium in a second, dilute layer, on top of those cesium atoms already chemisorbed. In migrating over the surface the atoms in this second layer could fill up any sites available in the first layer, or else return to the vapor phase. TAYLOR and LANGMUIR actually observed the formation of a second layer of cesium once a monolayer had been completed on the tungsten surface, and they identified this with the precursor state active at lower coverages. MOORE and ALLISON[5] have only recently found a similar constancy in the rate of adsorption of strontium and barium on tungsten, and have revived LANGMUIR's interpretation.

The evidence for the existence of a precursor and its effects on the rate of chemisorption of molecules is of the same character: for oxygen on tungsten (MORRISON and ROBERTS[2]) the sticking coefficient was found to decrease less rapidly, on increasing the amount chemisorbed, than would be predicted from the original Langmuir model requiring the collision of molecules from the gas with an empty site. MORRISON and ROBERTS[6] discovered a variation with oxygen pressure in the thermal accommodation coefficient of neon, on an oxygen-covered wire, and interpreted this to indicate two types of adsorption—a tightly bound composite layer stable to at least 1100°K, and a layer of molecular oxygen adsorbed on top of this, varying in concentration with pressure, and acting as a reservoir from which empty sites on the bare tungsten surface could be filled. The minimum lifetime of molecules in this state was estimated at two seconds, and the isosteric heat was later found to be 3000 calories (MORRISON and GRUMMITT[7]). The argument for the existence of a precursor presented by MORRISON and ROBERTS can be convincing, however, only if the sticking coefficient in the limit $\theta \to 0$, that is when a negligible fraction of the total surface is covered by chemisorbed material, is unity, as they indeed report; otherwise, a less rapid decrease of the sticking coefficient s with θ than expected from a simple $(1-\theta)$ dependence can be rationalized as well by assuming an activation energy for chemisorption decreasing continuously with θ. More recent work has not substantiated the value unity for the sticking coefficient, and it would seem that the stability of the layer identified by MORRISON and ROBERTS as the precursor is such that the dependence upon θ should have been significantly less than observed. We must conclude that this work does not constitute decisive evidence for the participation of a precursor state in chemisorption of molecules. The only significant rate data for the adsorption of molecules on a metal appear to be those for nitrogen on tungsten, obtained by BECKER and HARTMAN[3] and we shall concentrate our attention on this system.

The dependence of rate in this system upon the amount chemisorbed is somewhat similar to that reported for oxygen—the sticking coefficient s is constant (at 0·55 for $T = 300°\mathrm{K}$) with increasing θ, but decreases rapidly once a critical coverage θ_c (which at $T = 300°\mathrm{K}$ corresponds to $1{\cdot}25 \cdot 10^{14}$ molecules/cm^2) has been achieved. This constancy, as already mentioned, might be explained by a decrease in the activation barrier to chemisorption. However, the rate of adsorption at a fixed θ was found to decrease with increasing temperature; this immediately eliminates a simple activated adsorption, and suggests an intermediate step adversely affected by a temperature increase. The concept of a precursor adsorbed on the surface, with a concentration diminishing with increasing temperature, is intuitively appealing. Such a picture is by no means *necessarily* the one dictated by the decrease of rate with temperature, however.

If we assume that the overall rate of chemisorption per unit area (dn_2/dt) is limited by the reaction of a gas molecule (of mass M) and a single surface site,† then the rate equation, couched in the terminology of the transition state method, is

$$\frac{dn_2}{dt} = N_s\left[\frac{N_g kT}{V}\right]\frac{\mathcal{K}}{hf_z}\exp\left(-\frac{\Delta G^*}{kT}\right). \qquad (1)$$

N_s is the number of lattice sites available in unit area, N_g the number of gas molecules in volume V. The transmission coefficient for systems passing through the activated state is given by $\mathcal{K}$, f_z denotes the contribution to the partition function of

† This is presumably not appropriate for N_2 and is assumed only to simplify the presentation of the equations which follow.

the free molecule due to motion in a direction perpendicular to the surface, and ΔG^* is the free energy of activation for the transfer of a gas molecule, confined initially to a plane parallel to the surface, to the transition state. k, T and h are, as usual, the Boltzmann constant, absolute temperature, and PLANCK's constant, respectively. Substituting $(2\pi MkT/h^2)^{\frac{1}{2}}$ for the partition function f_z, and making use of the perfect gas laws, equation (1) becomes

$$s = \frac{1}{k_1 p}\frac{dn_2}{dt} = N_s \mathcal{K} \exp\left(\frac{-\Delta G^*}{kT}\right)$$
$$k_1 = (2\pi MkT)^{-\frac{1}{2}} \qquad (2)$$

and the dependence of rate upon temperature, keeping the number of chemisorbed molecules (n_2) constant, is given by

$$\frac{\partial \ln s}{\partial \frac{1}{kT}} = \frac{\partial \ln N_s}{\partial \frac{1}{kT}} + \frac{\partial \ln \mathcal{K}}{\partial \frac{1}{kT}} - \Delta H^*. \qquad (3)$$

The experimental results obtained for N_2 on tungsten demand only that the sum of the terms on the right-hand side of (3) must be greater than zero; this can be realized if with increasing temperature the number of adsorption sites N_s or the transmission coefficient $\mathcal{K}$ decreases, if the heat of activation is negative, or by a combination of these.

We can eliminate the first term as responsible for the decrease in rate; disordering of the low index faces of metal crystals through thermal fluctuations can occur only near the melting point, if at all (BURTON, CABRERA, and FRANK[8]). Moreover, studies in the field-emission microscope indicate that the activation energy for surface self-diffusion of the refractory metals is high, at least 1/3 the heat of sublimation, and that such diffusion on a tungsten surface is negligible below 800°K (MÜLLER[9]). Estimates of the transmission coefficient in chemisorption have been made by STRACHAN[10] and LENNARD-JONES[11], and these also suggest that no significant temperature dependence should arise due to this factor.

Granting that the first two terms in (3) do not suffice to establish a positive temperature coefficient as observed, and that $\partial \ln s/(1/kT) \sim -\Delta H^*$, this does not necessarily indicate an intermediate state prior to chemisorption; such a negative term could, for example, arise if the impinging molecule in the transition state has only limited lateral mobility.

Participation of a precursor to chemisorption *can* be established, however, from the very nature of the interactions between a gas molecule and a metal. As already pointed out by LENNARD-JONES[12], the potential energy of a molecule as a function of its separation from the surface shows two minima—a rather shallow one due to dispersion forces (less than 5 kcal per mole), at a distance comparable to the hard sphere diameter of the gas, and a very much more pronounced one close to the surface; the first energy trough corresponds to physical adsorption, the deep trough, in the immediate vicinity of the surface, to chemisorption. Molecules that upon colliding with the surface condense in this shallow energy trough can constitute the precursor, since they have a finite lifetime during which migration over the surface is possible (BRUNAUER[13]). To establish that such physically adsorbed molecules do participate in chemisorption, it is necessary only to demonstrate that the condensation of these molecules is not the limiting factor in the overall rate.

There are, at present, no quantitative studies of the condensation coefficient of molecules physically adsorbed on a surface, but indirect evidence indicates that energy transfer to the solid is efficient and that condensation occurs on impact. Specular reflection of molecular beams incident on a cleaved crystal face has been found only for hydrogen and helium atoms and hydrogen molecules (FRASER[14]). With other impinging species, adsorption and random evaporation has been held responsible for the diffuse scattering found. LENNARD-JONES and DEVONSHIRE[15] have estimated the condensation coefficient for physically adsorbed gas on a surface consisting of Debye oscillators, assuming that energy transfer between lattice and molecule occurs in one step. For a hydrogen molecule on a solid with a characteristic temperature of 300°($W \sim 310°$) their calculation yields a condensation coefficient, α, of 0·3. However, α increases both with temperature as well as with the mass of the incident molecule, so that for nitrogen a condensation coefficient close to unity is indicated. An increase in α with increasing temperature is to be expected at thermal

equilibrium from considerations of detailed balancing. From LENNARD-JONES and DEVONSHIRE it also follows that for an incident molecule with a fixed energy normal to the surface, corresponding to the experimental conditions under which nitrogen adsorption was measured (constant gas temperature, varying surface temperatures), the probability of condensation increases with increasing surface temperature.

Experimentally, the sticking coefficient for nitrogen on tungsten at room temperature was found to be 0·55, and to decrease with increasing temperature. In as much as rough estimates place the condensation coefficient close to unity, and an increase in temperature would increase α, it follows that a reservoir of physically adsorbed molecules must exist on the surface, and that these must be able to take part in chemisorption. It can be argued quite generally that any chemisorption reaction for which the conditions $s < 1, \alpha = 1$ are met must necessarily involve the participation of physically adsorbed precursors. In the following discussion we shall assume that the mechanism of chemisorption involves only physically adsorbed intermediates, and that the transmission coefficient for chemisorption is indeed unity; the justification of these simplifying assumptions must await further experiments, however.

3. THE RATE OF CHEMISORPTION

1. *Surface Processes Involving Precursor*

In view of the preceding, we will, in developing the kinetic formalism for chemisorption, consider only the processes that occur in the precursor state; we restrict ourselves to temperatures at which the precursor concentration (n_1 molecules/cm^2) is small compared with the number of atom sites N_s, and thus, because of the weak interactions with the surface, does not significantly affect the thermodynamic properties of the surface. Molecules that impinge upon the surface are presumed to be physically adsorbed; the barrier to lateral migration over the surface is small compared with that to evaporation, and in this state of binding, molecules can move over the surface to provide a reservoir from which chemisorption occurs. For the precursor state, we must account for evaporation from the surface with a rate n_1k_{-1}, transfer to the chemisorbed state, n_1k_2, and growth of the population itself, dn_1/dt; throughout we shall assume that evaporation of chemisorbed material can be neglected. The condition of mass balance establishes that the sum of these rates must equal the rate of supply of molecules to the surface, in this case to pk_1, the rate of impingement from the gas. Assuming a steady state for the precursor, ($dn_1/dt \sim 0$) we find for the sticking coefficient

$$s = \frac{1}{pk_1}\frac{dn_2}{dt} = 1 - \frac{k_{-1}}{k_{-1}+k_2} \qquad (4)$$

as shown previously (EHRLICH[16]). Except for the assumptions involving the precursor, this is quite general, and does not assume anything about the rate determining step in chemisorption. The constants k_i, however, denote the specific rate for unit area of surface, and incorporate terms for the number of sites available on the substrate. The interaction of molecules with a surface through dispersion forces is such that every surface atom will contribute to the adsorption of the precursor. For chemisorption, however, various structural factors have at one time or another been emphasized as important in affecting the rate. Qualitatively, an enhancement of reactivity at surface features such as crystal edges, grain boundaries, dislocations, etc., is clearly to be expected and we shall, therefore, elaborate (4) for several idealized models, both for a surface which is initially homogeneous and for one in which sites of different reactivity are present. For the latter, we will consider two limiting steps—limitation by the rate of diffusion of adsorbed precursor molecules to available reaction sites, and limitation by the rate of reaction between site and molecule. A third possibility, control through diffusion of chemisorbed species away from adsorption sites, will not be considered, inasmuch as the activation energy for such a step is orders of magnitude greater than that found for the system to be considered.

2. *Reaction Controlled Kinetics on a Homogeneous Surface*

This model has been previously discussed (EHRLICH[16]) and we shall limit ourselves to a consideration of the changes introduced by application of the transition-state method to the formulation of the rate equations. Using the nomenclature described above, the rate constants k_{-1} for the evaporation of physically-adsorbed species and k_2

VOL. 1 956/5

for the physically adsorbed→chemisorbed transition, are, respectively

$$k_{-1} = \frac{kT}{hf_z} \exp\left(\frac{-\Delta G^*_{-1}}{kT}\right) \tag{5}$$

$$k_2 = N_s \frac{kT}{hf_z} \exp\left(\frac{-\Delta G_2^*}{kT}\right). \tag{6}$$

from which it follows that

$$\frac{s}{1-s} = N_s \exp\left(\frac{\Delta G_{12}^*}{kT}\right) \tag{7}$$

$$\Delta G_{12}^* = \Delta G_{-1}^* - \Delta G_2^*. \tag{8}$$

In principle, we can estimate ΔG_{12}^*;† assuming that in the transition state the systems are ideal, and that the activated complex for chemisorption can be depicted as a molecule without rotational freedom and confined to a plane, while the activated complex for desorption of precursor can be approximated by a molecule in a twodimensional gas, we find (HILL[17])

$$\frac{s}{1-s} = \frac{N_s I_0^2(u) u \exp(-2u)}{[1-\exp(-Ku)]^2} \frac{\Theta_r h^2}{MkT^2} \exp\left(\frac{\epsilon_{-1}-\epsilon_2}{kT}\right) \tag{9}$$

$$u = \frac{V}{2kT} \qquad K = \left(\frac{2h^2}{Md^2V}\right)^{\frac{1}{2}}. \tag{10}$$

ϵ_{-1} and ϵ_2 are the zero-point energies of the activated complexes, $2V$ the maximum potential barrier to migration of the chemisorption complex over the surface, d the distance between potential minima on the surface (assuming a simple cubic surface lattice), Θ_r the characteristic temperature for molecular rotation and $I_0(u)$ is the modified Bessel function of the first kind of order zero. The temperature variation of rate over the range in which the complex can migrate over the surface (u small), reduces to

$$\frac{\partial \ln \dfrac{s}{1-s}}{\partial \dfrac{1}{kT}} = \epsilon_{-1} - \epsilon_2. \tag{11}$$

In the opposite extreme of an immobile transition state (u large)

$$\frac{\partial \ln \dfrac{s}{1-s}}{\partial \dfrac{1}{kT}} = 2kT + \epsilon_{-1} - \epsilon_2. \tag{12}$$

The mobile-immobile transformation is a continuous one and, according to the estimates of HILL, starts at $u \sim 16$ for chemisorbed gases, in good agreement with the experimental results of GOMER.[18] Such a transition can, on this model, result in a nonlinear dependence of rate upon temperature, but only over the limited temperature range of an immobile complex. TRAPNELL'S[19] attempt to explain the decrease of the rate of chemisorption with increasing temperature on this basis appears ill-advised.

From (9) we can estimate the sticking coefficient, assuming $N_s \sim 10^{15}$, a barrier to the lateral motion of the activated complex of ~ 30 kcal, and evaluating the potentials ϵ_{-1} and ϵ_2 from the temperature dependence. At 300°K, $s \sim 0{\cdot}04$, an order of magnitude lower than the experimentally determined one. For a transition state with rotational freedom within the plane of the surface, however, $s/(1-s)$ can be found from (9) by multiplying by $(\pi T/\Theta_r)^{\frac{1}{2}}$, and the coefficient of the first term in (12) is changed from 2 to 3/2. The calculated sticking coefficient, 0·4, is in excellent agreement with the experimental value. Though because of the drastic assumptions this agreement appears fortuitous, it is clear that the entropy of activation for chemisorption must be of the same order as that for desorption of precursor—no serious restriction upon the degrees of freedom of the precursor is permitted. It should be apparent that this more detailed presentation does not significantly alter the conclusions previously reached (EHRLICH[16]). Also, it follows from (7) that although s may be insensitive to changes in θ, the ratio $s/(1-s)$ should, for this model, vary linearly

† It should be apparent that ΔG^* is a standard quantity and independent of concentration only for ideal systems. For a concentrated system, such as the mixture of filled and empty sites, this approximation may become invalid; ΔG^* must then include activity corrections, and it becomes a concentration-dependent quantity. The precursor, however, is, on our model, present in only small amounts; and deviations from ideality for it will therefore be negligible.

with the fraction of the total area covered by chemisorption, unless we invoke a concentration dependence for ΔG_{12}^*.

3. *Diffusion-Controlled Kinetics on a Heterogeneous Surface*

We now consider the opposite situation—we shall assume that every molecule that during its meandering over the surface finds an appropriate site is immediately chemisorbed. Such a model has to be considered only if every atom site on the surface does not constitute an adsorption site, that is, for a heterogeneous surface. For a surface in which this is not the case, the rate of impingement would, in the limit $\theta \to 0$, be controlling and $s = 1$. We shall discuss two idealized surfaces—one in which the a priori unavailable sites, constituting a fraction A_B of the total number of atom sites, are randomly distributed, and another surface, for which the unavailable sites are clustered together, leaving patches amounting to A_A of the total area, on which adsorption *can* occur. For the first of these surface configurations, the steady state condition requires that $k_{-1}/(k_{-1}+k_2)$ be equal to $(A_A\theta+A_B)^{m+1}$, m being the number of atom sites visited by a physically adsorbed molecule (on a surface covered with a monolayer, $\theta = 1$), and θ the fraction of the A sites already reacted. That is, evaporation can occur only if a precursor molecule does not at any time during its sojourn on the surface collide with an available site. As shown by MORRISON and ROBERTS[2], relation (4) becomes

$$s = 1-(A_A\theta+A_B)^{m+1}. \tag{13}$$

m, the number of sites visited during random movement over the surface, is given by BURTON, CABRERA, and FRANK[8] as

$$m = \exp\frac{(\Delta H_{-1}^*-\Delta H_s^*)}{kT} \tag{14}$$

provided the entropy of activation for evaporation and surface diffusion is the same. ΔH_s^* is the activation energy for surface diffusion, ΔH_{-1}^* that for evaporation of physically adsorbed molecules from sites which have already undergone chemisorption. The temperature coefficient of the rate of chemisorption, assuming $m \gg 1$, is most conveniently given by

$$\frac{\partial \ln\dfrac{s}{1-s}}{\dfrac{1}{\partial kT}} = (\Delta H_{-1}^*-\Delta H_s^*)\left(1-\frac{s}{2}+\frac{s^2}{3}\ldots\right). \tag{15}$$

The sticking coefficient, under the assumptions of this model, should differ significantly from unity only when the fraction of sites either intrinsically unavailable, or occupied, approaches unity. For an initially clean surface ($\theta \to 0$), $s < 1$ can be achieved only if $A_B{}^{m+1}$ is comparable to unity itself, and, just as above, this means that the sticking coefficient will decrease with increasing adsorption (θ) unless m likewise increases.

A situation formally similar to this model would exist if chemisorption can occur only after a physically adsorbed molecule has been activated† at some singularity, such as an emergent screw dislocation, or an impurity center. Simplifying the problem by assuming again that every activated molecule can find a permanent adsorption site (which would be true in the limit as $\theta \to 0$) and that the activation centers are randomly distributed,

$$s = 1-A_D{}^{m+1} \tag{16}$$

where A_D constitutes the fraction of the sites that are nonactivating. Equations (14) and (15) still apply, of course.

For the second model, that of a patch surface, we assume that a molecule impinging upon an unavailable B patch will be adsorbed only if its peregrinations bring it across the boundaries and into a region of available A sites. We idealize the surface as consisting of circular patches of unavailable sites, with radius y_0, surrounded by regions in which chemisorption can occur, and assume that a molecule will be able to escape from the former if its point of arrival is within a distance x_s of the edge. The ratio of molecules that are adsorbed to the number impinging per unit time is thus, in the limit $\theta \to 0$, just equal to the ratio of the

† The need for considering certain singularities in the surface as activation centers only arises from the fact that experimentally the number of molecules that can be adsorbed has been found to be of the same order of magnitude as the number of lattice sites. Any model that assumes that all adsorption occurs on singularities must therefore fail.

VOL. 1 1956/57

collecting area to the total area, that is, neglecting terms in x_s^2

$$\lim_{\theta \to 0} s = A_A\left(1+\frac{A_B 2x_s}{A_A y_0}\right). \tag{17}$$

The first term accounts for molecules that are adsorbed after impinging on the good sites, the second term for those adsorbed after escaping from B patches. For higher coverages, this expression must be modified by multiplying by a factor of the form of (13), i.e., $(1-\theta^m)$; this is an approximation, since it denies the possibility that a physically adsorbed molecule can escape an A patch by surface diffusion; this will be considered more exactly later. If we substitute for x_s the root mean square separation between the point of arrival and departure of a physically adsorbed molecule,

$$\frac{x_s}{d} = \exp\left(\frac{\Delta H_{-1}^* - \Delta H_s^*}{2kT}\right) \tag{18}$$

we find for the sticking coefficient the expression

$$s = A_A\left[1+\frac{A_B 2d}{A_A y_0}\exp\left(\frac{\Delta H_{-1}^* - \Delta H_s^*}{2kT}\right)\right]. \tag{19}$$

The temperature dependence is now most significantly obtained from

$$\frac{\partial \ln\left(\frac{s}{A_A}-1\right)}{\partial \frac{1}{kT}} = \frac{\Delta H_{-1}^* - \Delta H_s^*}{2}. \tag{20}$$

As already indicated, the dependence upon the amount of chemisorption will be of the order of $(1-\theta^m)$. The presence of sites that are intrinsically unavailable is accounted for by a multiplicative factor, and therefore, we can, on this basis, reconcile a sticking coefficient of less than unity, with a lack of variation of s upon increasing θ without invoking a dependence of the activation terms upon coverage.

A trivial modification of equation (17) describes the rate if we assume that chemisorption occurs after physically adsorbed molecules are activated at the boundaries of grains of radius y_0, and that the diffusion from the point of impingement to the grain boundary is limiting. For this case we can equate to zero the fraction of the total area occupied by boundaries, and with this simplification, equations (19) and (20) become

$$s = \frac{2d}{y_0}\exp\left(\frac{\Delta H_{-1}^* - \Delta H_s^*}{2kT}\right) \tag{21a}$$

$$\frac{\partial \ln s}{\partial \frac{1}{kT}} = \frac{\Delta H_{-1}^* - \Delta H_s^*}{2}. \tag{21b}$$

4. *Reaction Controlled Kinetics on a Heterogeneous Surface*

The possibility of the escape of precursor molecules from an active patch is most readily treated for a model in which the reactive sites are formed into parallel strips (A)† of width $2a$, extending over the entire area and separated from one another by an inactive region (B) of width $2y_0$. Precursor molecules are supplied to a reactive strip (A) both directly from the gas and also by surface diffusion. From the steady state condition on the concentration of precursor it follows that the net diffusive flux into unit length of A, that is, the difference in the rates at which molecules jump from B into A, $n_{1B}k_{BA}$, and from A into B, $n_{1A}k_{AB}$, must equal the difference between the rate of impingement and evaporation for the adjoining strip of B extending from the midpoint of the region B to the boundary with A. The steady-state condition also requires that this net diffusive flux be equal to the difference between the rate at which precursor is eliminated in strip A by reaction and evaporation, and the rate at which it is supplied by direct impingement. These requirements are given respectively, by

$$n_{1B}(y_0)k_{BA} - n_{1A}k_{AB} = -k_{-1B}\int_0^{y_0}(n_{1B}-n_{1B}^{\circ})dy \tag{22a}$$

$$n_{1B}(y_0)k_{BA} - n_{1A}k_{AB} = n_{1A}a(k_{-1A}+k_2) - pk_1a \tag{22b}$$

† This development follows closely the work of BURTON and CABRERA.[8] A formal adaptation of their results to the discussion of surface catalytic reactions has been presented by PATTERSON.[20]

n_{1B}° indicates the concentration of precursor in equilibrium with the gas at the given pressure, and $n_{1B}(y)$ the actual concentration in the region B as a function of the distance y from the locus of midpoints of B. The term $(n_{1B}-n_{1B}^{\circ})$ can be found from the equation for surface diffusion

$$(n_{1B}-n_{1B}^{\circ}) = x_s^2 \nabla^2 n_{1B} \tag{23}$$

which, for an array of such parallel strips has the solution

$$(n_{1B}-n_{1B}^{\circ}) = Y_0 \cosh(y/x_s). \tag{24}$$

Solving (22a), (22b), and (24) for the concentration of the precursor in strip A, n_{1A}, the sticking coefficient, $s = \dfrac{a}{y_0+a}\dfrac{n_{1A}k_2}{pk_1}$, is obtained as

$$s = A_A \frac{\left[1+x_s\frac{k_{-1B}}{k_{BA}}\tanh\left(\frac{y_0}{x_s}\right)+\frac{x_s}{a}\tanh\left(\frac{y_0}{x_s}\right)\right]k_2}{k_2\left[1+x_s\frac{k_{-1B}}{k_{BA}}\tanh\left(\frac{y_0}{x_s}\right)\right]+k_{-1A}\left[1+x_s\frac{k_{-1B}}{k_{BA}}\tanh\left(\frac{y_0}{x_s}\right)\right]+x_s\frac{k_{-1B}}{k_{BA}}\frac{k_{AB}}{a}\tanh\left(\frac{y_0}{x_s}\right)}. \tag{25}$$

The rate constants entering into (25) have all been discussed above, with the exception of k_{BA} and k_{AB} which are, respectively (POUND, SIMNAD, and YANG[21]).

$$k_{BA} = \frac{d}{4}\frac{kT}{hf_z}\exp\left(\frac{-\Delta G_{S_B}^{*}}{kT}\right) \tag{26a}$$

$$k_{AB} = \frac{d}{4}\frac{kT}{hf_z}\exp\left(\frac{-\Delta G_{S_A}^{*}}{kT}\right). \tag{26b}$$

Using (5), (18), and (26a) the factor $x_s k_{-1B}/k_{BA}$ appearing in (25) can be simplified to $4dx_s^{-1}$. Three distinct temperature ranges can be distinguished; for low enough T, $x_s \gg y_0$; in an intermediate temperature range $y_0^2 \gg x_s^2 \gg 4da$, and at high temperatures $a \gg x_s$. For the region of low temperatures, $1 \gg 4dx_s^{-2}y_0$, and the sticking coefficient becomes

$$s = \frac{k_2}{k_2+k_{-1A}+4dx_s^{-2}y_0(k_{AB}/a)}. \tag{27}$$

This, except for the last term in the denominator, is identical with the expression for a uniform surface (EHRLICH[16]). At low enough temperatures, however, $k_{-1A} < 4dx_s^{-2}(y_0/a)k_{AB}$; i.e., precursor may be lost more rapidly by diffusion than by direct evaporation. Under these conditions, every molecule that impinges upon the surface will encounter an active A strip, and the rate of chemisorption is determined solely by the relative magnitude of the rate constants involved. For $k_2 \gg (k_{AB}/a) \gg k_{-1A}$, the sticking coefficient will be unity, and insensitive to increasing θ, the fraction of the active A sites already reacted. This condition can be met only if $\Delta H_2^* < \Delta H_{-1}^*$. At higher temperatures, at which still $1 \gg 4dx_s^{-1}$

$$s = A_A\frac{(1+x_s/a)k_2}{k_2+k_{-1A}+4dx_s^{-1}(k_{AB}/a)}. \tag{28}$$

The area from which gas molecules can reach the reactive region is now limited to a strip of width x_s, as already indicated above. For $k_2 \gg k_{-1A}$, $k_2 \gg (k_{AB}/a)$ the overall rate becomes diffusion controlled, and the sticking coefficient is of the form of (17). The temperature dependence in this range is given by

$$\frac{\partial \ln s}{\partial \frac{1}{kT}} = \frac{\Delta H_{-1}^{*}-\Delta H_s^{*}}{2}\left(1-\frac{a}{x_s}..\right). \tag{29}$$

Under these circumstances, s can be smaller than unity and still independent of θ. The range of θ $(0 < \theta < \theta_c)$ for which the rate constants will satisfy the condition $k_2 \gg k_{-1A}$, $k_2 \gg k_{AB}/a$, depends upon the temperature as indicated by (5), (6), and (26b) and for $\Delta H_2^* < \Delta H_{-1}^*$ will decrease as the temperature is increased. If the activation terms are not so related, a sticking coefficient independent of θ cannot be expected.

At still higher temperatures, $a > x_s$ and

$$s = A_A\frac{k_2}{k_2+k_{-1A}+\frac{k_{AB}}{a}\left(\frac{4d}{x_s+4d}\right)}. \tag{30}$$

Now only the direct transport from the gas is

significant. Also at higher temperatures, if $\Delta H_2^* < \Delta H_{-1}^*$, the condition $k_2 \gg k_{-1A}$, $k_2 \gg (k_{AB}/a)$ will no longer be met, and the sticking coefficient will decrease with increasing θ. Inasmuch as ΔH_s^*, the heat of activation for surface diffusion, is small compared with ΔH_{-1}^*, the heat of activation for evaporation of physically adsorbed molecules, the temperature dependence will approach the limit

$$\frac{\partial \ln s}{\partial \frac{1}{kT}} = \Delta H_{-1}^* - \Delta H_2^*. \tag{31}$$

This is a specific illustration of the general possibility of a change in the rate controlling step of chemisorption. If $\Delta H_s^* - \Delta H_2^* > 0$, then even though at a low temperature the diffusion to an adsorption site is limiting, the reaction at the site may be controlling at more elevated temperatures.

4. THE MECHANISM OF CHEMISORPTION—COMPARISON WITH EXPERIMENT

Two features are striking about the rate of chemisorption of nitrogen on tungsten, as determined by BECKER and HARTMAN[3]—the decrease of the rate at higher temperatures, and the constancy of the sticking coefficient for increasing adsorption. Qualitatively, this behavior can be understood in terms of any of the models presented above—the presence of a precursor on the surface is in accord with the variation of rate with temperature, and would obviously also bring about a less pronounced dependence of the sticking coefficient upon θ. A quantitative comparison with experiment permits us to eliminate all but one of these, however.

We can dispose of the possibility that isolated dislocations or impurities, through their local disturbance of the surface dipole layer of the metal, contribute significantly to the overall rate by appealing to equation (16). From the plot of rate of chemisorption against temperature, shown in Fig. 1, we can estimate $\Delta H_{-1}^* - \Delta H_s^*$ at $\sim 2{\cdot}2$ kcal, that is the number of sites visited by the precursor at 300°K prior to evaporation would be ~ 40. The maximum number of dislocations or impurities emerging on a tungsten surface has been recently estimated as one per thousand atoms (DRECHSLER, PANKOW, and VANSELOW[22]). Even assuming so large a number (approximately that for heavily cold-worked materials), and neglecting that every collision with such a singularity might not be effective, we find, at 300°K, a sticking coefficient, s of only $\sim 0{\cdot}04$. The alternative activation mechanism—due to grain boundaries—likewise can be shown to disagree with experience.

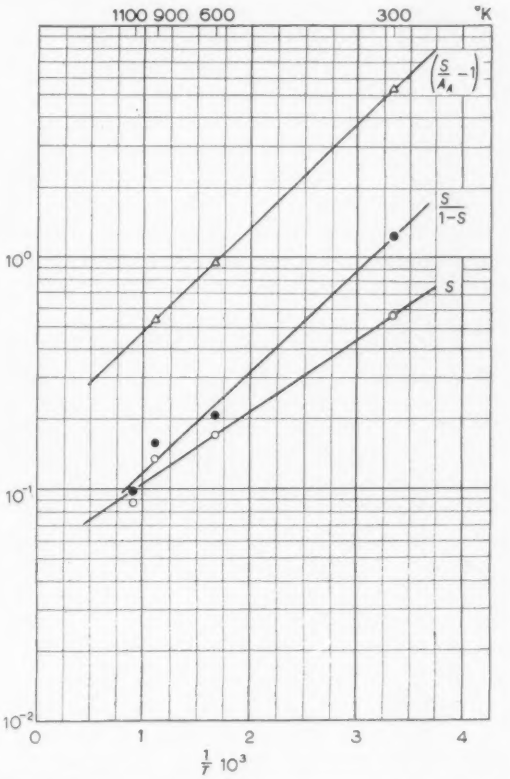


FIG. 1. The variation with temperature of the rate of chemisorption (BECKER and HARTMAN[3]), of nitrogen on tungsten at $\theta \to 0$. Plot according to:

Equation (17) △
Equation (15) .
Equation (29) o

From the temperature dependence, interpreted according to equation (21b), we find x_s at 300°K to be ~ 10 atom spacings, and this would require grains with a radius of only ~ 40 atom spacings in order to agree with the experimentally determined sticking coefficient. Such grains are at least five orders of magnitude smaller than actually observed (BECKER[23]).

As indicated in the previous section, reaction-controlled kinetics on an initially homogeneous crystal surface can yield a value of $s/(1-s)$ independent of θ, as shown by experiment, only if the free energy of activation, ΔG_{12}^*, increases as n_2 increases. Such an increase cannot be attributed to the heat term ΔH_{12}^*, however, as this would bring about a constant $s/(1-s)$ ratio at one temperature only. A dependence of the entropy of activation upon θ, besides being difficult to envisage for a uniform surface, would still leave unexplained the decrease of θ_c with increasing temperature, and we must therefore conclude that a uniform surface model cannot be brought into agreement with the kinetics observed experimentally. The same considerations, applied to diffusion controlled kinetics on a surface with randomly distributed imperfections, will eliminate the second model considered.

For a patchy surface, if we accept the assumption that in the limit $\theta \to 0$ every molecule that encounters a good patch is chemisorbed, we can evaluate A_A from the sticking coefficient in the limit $T \to \infty$, that is when surface transport makes a negligible contribution. The value of $\Delta H_{-1}^* - \Delta H_s^*$ obtained from (20) is 3·8 kcal, giving a mean distance traversed on the surface, x_s, of 24 atom diameters at 300°K. This leads to unavailable patches having a radius of ~ 100 atom spacings and occupying 0·92 of the surface. There are obvious difficulties with this simple picture. If we assume $\sim 10^{15}$ atom sites per cm^2 of macroscopic surface, we would expect that only $\sim 50 \cdot 10^{12}$ molecules should be adsorbed. This is contrary to observation, but it is possible to rationalize this by assuming that the surface is, on a microscopic scale, rough, so that the actual area exceeds the apparent area computed from the bulk dimensions. A discrepancy of an order of magnitude would appear to be too large to be accounted for in this way. However, the division of the surface into only two kinds of patches—good and bad—is, of course, artificial, and if we replace this by a more elaborate distribution of patches of different reactivity, with only the most reactive ones adsorbing initially, this difficulty is relieved. Above $T = 600$°K, however $(ds/d\theta)_{\theta=0}$ is not equal to zero, that is the condition that every molecule striking a good patch be adsorbed does not seem to be met. Once we relax this stringent condition, we have the more comprehensive model developed in Section 3. 4. In the low temperature region ($T < 900$°K), the condition $k_2 \gg k_{-1A}$, $k_2 \gg (k_{AB}/a)$ is evidently satisfied, and the decrease in sticking coefficient s with increasing temperature is due to the decrease in surface transport, as indicated by (28). For 900°K and higher, the limiting step becomes the rate of reaction—even for $\theta \to 0$ the evaporation and diffusion of the physically-adsorbed molecules dominates. By resorting to equation (29), we obtain for $\Delta H_{-1}^* - \Delta H_s^*$ an approximate value of 3 kcal, and assuming $\Delta H_s^* \sim \frac{1}{3} \Delta H_{-1}^*$, we find $\Delta H_{-1}^* \sim 4{\cdot}5$ kcal. The data at higher temperatures is not sufficient for a quantitative determination of $(\Delta H_{-1}^* - \Delta H_2^*)$ but this appears to be of the order of 2 kcal. The heat of desorption of the precursor, deduced on this model, is of the correct magnitude for dispersion forces, consistent with the initial assumption of physically-adsorbed molecules acting as precursor. Also, the identification of the temperature dependence of the sticking coefficient, for $T < 900$°K, with $\Delta H_{-1}^* - \Delta H_s^*$, plus the requirement $\Delta H_{-1}^* < \Delta H_2^*$, permit us to conclude that the activation energy for chemisorption is small, at most of the order of 1 or 2 kcal.

Though the onset of reaction control prevents a straightforward evaluation of A_A from the high-temperature data, this can be estimated roughly, from the value of the sticking coefficient for which $(ds/d\theta)_{\theta=0} \neq 0$, as $\sim 1/9$. This yields for a, the halfwidth of a reactive step, a value of 3 atom diameters, and for $y_0 \sim 21$. These estimates are admittedly crude, but the order of magnitude is clearly dictated by the importance of the diffusive transport at temperatures as high as 600°K. Such active patches of essentially atomic dimensions can be associated with only one feature of a solid surface-lattice steps formed during growth or evaporation. An a priori calculation of the extent to which steps would occur on tungsten ribbon is not feasible because of the uncontrolled conditions under which the surfaces are formed. That lattice steps can contribute significantly to the total surface area is evident, however, from field ion microscope pictures of tungsten single crystals (DRECHSLER *et al*[22]) which indicate lattice steps separated from one another by only 25Å—the surface appears to be made up of terraces of low index faces. With such a spacing, for example, we have

VOL. 1 1956/5

4×10^6 cm of steps per cm^2 of surface, which, if 2·5Å high, would provide an area of 0·1cm^2.

The question arises whether the high reactivity of A patches is to be associated with lattice steps per se, or whether this arises because atom planes of the correct arrangement happen to be subdivided in this fashion. An unequivocal answer cannot be inferred from the experimental results; one can say, however, that the dimensions of the active patches are such that the distortion of the surface dipole layer of the metal due to atomic roughness would be significant (HERRING[24]) and that, moreover, the appearance of a particular lattice plane in this state of subdivision is unlikely. An attempt to deduce an effect of lattice edges upon the reactivity of tungsten toward the adsorption of oxygen from observations in the field-emission microscope has recently been made by BECKER[23]. The interpretation so far offered of changes in emission patterns is not unique, however, and BECKER'S conclusion, though possibly correct, is not warranted by the evidence.

Fundamental to our analysis of the kinetics of adsorption and to the conclusion that surface heterogeneities of atomic dimensions play a dominant rôle, is the initial assumption that physically-adsorbed molecules constitute the precursor to chemisorption; it is on this basis, for example, that the steady state assumption ($dn_1/dt = 0$) is justified, or the elimination of diffusion away from a site as a limiting mechanism. The interpretation of the results arrived at on these assumptions is indeed consistent with the starting premises. Moreover, although recent work (EHRLICH[25]) indicates that other states of binding, beyond the single chemisorbed and physically adsorbed level considered here, must be accounted for on the surface, preliminary results (EHRLICH[26], EHRLICH and HICKMOTT[27]) do point to a presursor physically adsorbed on the surface.

The kinetics of chemisorption have been most adequately explored for nitrogen on tungsten, but preliminary experiments also indicate a sticking coefficient significantly less than unity, yet initially independent of θ, for carbon monoxide (EHRLICH[28]) as well as oxygen on tungsten (BECKER and BRANDES[29]). It is these features which for nitrogen lead to the picture of preferential reaction at active patches, and a similar conclusion for these systems seems indicated.

Acknowledgement—Stimulating discussions with D. TURNBULL assisted in formulating some of the ideas presented here.

REFERENCES

1. TAYLOR J. B. and LANGMUIR I. *Phys. Rev.* **44,** 423 (1933).
2. MORRISON J. L. and ROBERTS J. K. *Proc. Roy. Soc.* **A173,** 13 (1939).
3. BECKER J. A. and HARTMAN C. D. *J. Phys. Chem.* **57,** 157 (1953).
4. LANGMUIR I. *Chem. Rev.* **6,** 451 (1929).
5. MOORE G. E. and ALLISON H. W. *J. Chem. Phys.* **23,** 1609 (1955).
6. MORRISON J. L. and ROBERTS J. K. *Proc. Roy. Soc.* **A173,** 1 (1939).
7. MORRISON J. L. and GRUMMIT W. E. *J. Chem. Phys.* **21,** 654 (1953).
8. BURTON W. K., CABRERA N. and FRANK F. C. *Phil. Trans. Roy. Soc.* **A243,** 299 (1951).
9. MÜLLER E. W. *Ergeb. exakt. Naturwiss.* **27,** 290 (1953).
10. STRACHAN C. *Proc. Roy. Soc.* **A158,** 191 (1937).
11. LENNARD-JONES J. E. and GOODWIN E. T. *Proc. Roy. Soc.* **A163,** 101 (1937).
12. LENNARD-JONES J. E. *Trans, Faraday Soc.* **28,** 333 (1932).
13. BRUNAUER S. *The Adsorption of Gases and Vapors,* p. 448, Princeton University, Princeton, N.Y. (1945).
14. FRASER R. G. J. *Molecular Rays,* chap. 3, Cambridge University Press, Cambridge, (1931).
15. LENNARD-JONES J. E. and DEVONSHIRE A. F. *Proc. Roy. Soc.* **A156,** 6 (1936).
16. EHRLICH G. *J. Phys. Chem.* **59,** 473 (1955).
17. HILL T. L. *J. Chem. Phys.* **14,** 441 (1946).
18. GOMER R. and WORTMAN R. *J. Chem. Phys.* **23,** 1741 (1955).
19. TRAPNELL B. M. W. *Chemisorption,* p. 25, Academic Press Inc., New York, N.Y. (1955).
20. PATTERSON D. *Trans. Faraday Soc.* **49,** 802 (1953).
21. POUND G. M., SIMNAD M. T. and YANG L. *J. Chem. Phys.* **22,** 1215 (1954).
22. DRECHSLER M., PANKOW G. and VANSELOW R. *Z. physik. Chem.* **4,** 249 (1955).
23. BECKER J. A. *Advances in Catalysis* **7,** 135 (1955).
24. HERRING C., in *Metal Interfaces,* American Society for Metals, Cleveland, Ohio, (1952).
25. EHRLICH G. *J. Chem. Phys.* **23,** 1543 (1955).
26. EHRLICH G. *J. Chem. Phys.* **24,** 482 (1956).
27. EHRLICH G. and HICKMOTT T. W. *Nature* **III,** 1045 (1956).
28. EHRLICH G. *J. Phys. Chem.* **60,** October (1956).
29. BECKER J. A. and BRANDES R. G. *J. Chem. Phys.* **23,** 1323 (1955).

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 14–22.

EFFECTIVE MASS APPROXIMATION FOR EXCITONS*

G. DRESSELHAUS†

Institute for the Study of Metals, University of Chicago, Chicago, Illinois

(*Received* 4 *May* 1956)

Abstract—The effective mass approximation for degenerate electronic energy bands is applied to the interaction of electrons and holes in an insulator. It is shown that bound exciton states result from this treatment. The exciton Hamiltonian is of the form derived previously for impurity-state problems. The criterion for the validity of the effective mass approximation is that the lowest state be bound with an energy $<0{\cdot}2$ eV. Corrections to the approximation to account for the effective magnetic moments of the electron and hole are indicated. Selection rules and line widths for optical excitation of excitons are discussed. The question of the existence of positronium in crystalline solids is also considered.

VOL.
1
1956/5

1. INTRODUCTION

IT was FRENKEL[1] who first proposed the existence of nonconducting excited electronic states in insulators. His original idea was that individual atoms in a solid could be excited and this excitation would travel in a particle-like fashion through the solid. These packets of excitation are termed excitons. The exciton concept has been used to explain certain optical absorption lines observed in pure solids.

Two distinctly different treatments for the exciton problem have been formulated, each with its own range of applicability. In ionic or molecular crystals an atomic orbital approach along the lines first proposed by FRENKEL[2] and extended recently by OVERHAUSER[3] and others[4] gives the more exact solution for the lowest bound exciton states. This treatment involves formulating exciton wave functions from atomic orbitals of the atoms at the crystal lattice sites. The orginal Frenkel proposal considered linear combinations of functions representing individual atoms in the crystal being excited (i.e. a combination of electron and hole at the same lattice site). The formation of functions representing an electron excited to an atomic orbital on a nearest-neighbor atom (electron and hole on nearest neighbors) has been discussed most recently by OVERHAUSER.[3] Further extensions to include electrons excited to next nearest neighbors, etc. may easily be made using the techniques outlined in Ref. (3). The atomic orbital approach, however, leads in general only to a qualitative picture of the overall exciton spectrum. For weakly bound exciton states the necessity of including large contributions from the highly excited atomic orbital functions (i.e. states representing electrons and holes widely separated) makes even this qualitative treatment impractical.

An alternate treatment using the effective mass approximation for the exciton problem was first given by WANNIER.[5] His proposal for an exciton can be most clearly pictured as an electron in a conduction band being bound to a hole in a valence band. The validity of the effective mass approximation depends on having the interaction potential slowly varying over the dimensions of a unit cell; i.e., the exciton dimension is large compared to the lattice constant. This formulation of the perturbed periodic potential has been used quite extensively for impurity-state problems in semiconductors.[6, 7]

Recently, absorption bands, which may correspond to excitons as treated by an effective mass approximation, have been seen in Cu_2O,[8,9] CdS,[10] HgI_2, PbI_2, and CdI_2.[11] Cu_2O, cubic in structure, exhibits a hydrogenic spectrum which indicates that the simple spherical energy band model would be applicable to the valence and conduction-band extrema. The energy surfaces at band extrema for nondegenerate positions in the Brillouin

* Supported in part by the U.S. Office of Naval Research and by the U.S. Air Force through a contract with Chicago Midway Laboratories.

† Present address: Department of Physics, Cornell University, Ithaca, New York.

zone, however, are in general elliptical in nature. The center of the zone in cubic structures is one of the few cases where spherical surfaces may occur. The possibility of degenerate extrema is also to be considered in the general formulation of an effective mass approximation for excitons. In addition to the gross features of the spectrum in Cu_2O, GROSS *et al.*[8] have observed fine structure and Zeeman splittings of a few of the exciton levels. Therefore, a discussion of the effects expected from the effective magnetic moments of the hole and electron forming the exciton is also required. With the development of these considerations in mind, one may then re-examine the optical absorption data in solids. The correct interpretation of exciton spectra may provide a useful tool in determining the electronic band structure in some insulators.

In this paper the effective mass approximation for excitons is extended to include degeneracies at the energy band extrema arbitrarily located within the Brillouin zone. The Hamiltonian which gives the exciton spectrum at the exciton-band extrema is derived, and the energy eigenvalues are discussed for a few examples. A perturbation theory treatment is formulated which gives the effective mass of the exciton band near the band minimum. This perturbation theory is similar to that for electronic energy bands.[12] A discussion of the failure of the effective mass formulation of the problem for tightly bound excitons, e.g. in ionic crystals,[3] is given. The selection rules for optical excitation of excitons are treated and reasons given for the absence of exciton absorption lines in certain semiconductors. Finally, the similarity between the exciton and positronium in solids is briefly considered.

2. DEVELOPMENT OF THE EXCITON HAMILTONIAN

In developing the effective mass theory for excitons we shall first treat the case of two spherical energy surfaces. Though the solution of this problem for nondegenerate energy surfaces is well known, the treatment given here is more readily extended to the more general case than is the usual formulation. In an effective mass approximation the Hamiltonian is:

$$H=-\frac{\hbar^2}{2m_e}\nabla_{r_e}{}^2-\frac{\hbar^2}{2m_h}\nabla_{r_h}{}^2-\frac{e^2}{\epsilon|\mathbf{r}_e-\mathbf{r}_h|}, \tag{1}$$

where m_e and m_h are the electron and hole effective masses, $\mathbf{r}_e$ and $\mathbf{r}_h$ are their respective co-ordinates, and ϵ is the dielectric constant of the crystal. Making the co-ordinate transformation*

$$\begin{aligned}\boldsymbol{R}&=\tfrac{1}{2}(\boldsymbol{r}_e+\boldsymbol{r}_h)\\ \boldsymbol{r}&=\boldsymbol{r}_e-\boldsymbol{r}_h,\end{aligned} \tag{2}$$

the Hamiltonian may be written

$$\begin{aligned}H=&-\frac{\hbar^2}{2\mu}\nabla_r{}^2-\frac{e^2}{\epsilon r}\\ &-\frac{\hbar^2}{2}\left(\frac{1}{m_e}-\frac{1}{m_h}\right)\nabla_r\cdot\nabla_R-\frac{\hbar^2}{2\mu}\nabla_R{}^2\end{aligned} \tag{3}$$

where $\mu=m_em_h(m_e+m_h)^{-1}$ is the reduced mass. The operator ∇_R commutes with the Hamiltonian, hence the eigenfunctions may be taken in the form

$$\Psi=e^{i\mathbf{K}\cdot\mathbf{R}}\psi(r), \tag{4}$$

where $\mathbf{K}$ is the wave vector for the exciton and satisfies the equation

$$\begin{aligned}\Big\{-\frac{\hbar^2}{2\mu}\nabla_r{}^2-\frac{e^2}{\epsilon r}-\frac{\hbar^2}{2}i\left(\frac{1}{m_e}-\frac{1}{m_h}\right)K\cdot\nabla_r+\\ +\frac{\hbar^2}{2\mu}K^2\Big\}\psi(r)=E\psi(r).\end{aligned} \tag{5}$$

The dispersion relation for the exciton band may now be found using perturbation theory for the $\mathbf{K}\cdot\nabla_r$ term. This procedure is completely analogous to the perturbation theory treatment[13] for electronic energy bands. The unperturbed Hamiltonian

$$H^\circ=-\frac{\hbar^2}{2\mu}\nabla_r{}^2-\frac{e^2}{\epsilon r} \tag{6}$$

has energy eigenvalues

$$E_n{}^\circ=-\frac{\mu e^4}{2\hbar^2\epsilon^2}\frac{1}{n^2} \tag{7}$$

and hydrogenic eigenfunctions ϕ_n. Thus the excitons at the center of the zone have a hydrogenic spectrum given by eq. (7). For small values of $\mathbf{K}$, a perturbation expansion to terms in $\mathbf{K}^2$ gives

* The Jacobian of the transformation is unity, hence, $\int F(\mathbf{r}_e,\mathbf{r}_h)\,d\mathbf{r}_e d\mathbf{r}_h=\int F(\tfrac{1}{2}\mathbf{r}+\mathbf{R},-\tfrac{1}{2}\mathbf{r}+\mathbf{R})\,d\mathbf{r}d\mathbf{R}$

$$E_n(\boldsymbol{K}) = E_n^{\circ} + \frac{\hbar^2 K^2}{2\mu} + \frac{\hbar^2}{4}\left(\frac{1}{m_e} - \frac{1}{m_h}\right)^2 \sum_{n'}{}' \frac{(\phi_n|\boldsymbol{K}\cdot\boldsymbol{p}|\phi_{n'})(\phi_{n'}|\boldsymbol{K}\cdot\boldsymbol{p}|\phi_n)}{E_n^{\circ} - E_{n'}^{\circ}}, \tag{8}$$

where $\boldsymbol{p} = -i\hbar\nabla_r$.
Using the f-sum rule,

$$\frac{2}{\mu}\sum_{n}{}' \frac{(\phi_n|p_\alpha|\phi_{n'})(\phi_{n'}|p_\beta|\phi_n)}{E_n^{\circ} - E_{n'}^{\circ}} = -\delta_{\alpha\beta}. \tag{9}$$

Eq. (8) yields

$$E_n(\boldsymbol{K}) = E_n^{\circ} + \tfrac{1}{2}\hbar^2(m_e + m_h)^{-1}K^2. \tag{10}$$

This result could have been obtained immediately if the transformation to the center of mass co-ordinates

$$\begin{aligned} \boldsymbol{R}_{C.M.} &= (m_e + m_h)^{-1}(m_e\boldsymbol{r}_e + m_h\boldsymbol{r}_h) \\ \boldsymbol{r} &= \boldsymbol{r}_e - \boldsymbol{r}_h \end{aligned} \tag{11}$$

had been made instead of the transformation (2). However, a center of mass transformation cannot be defined for excitons formed from degenerate bands. Eq. (10) shows that all excitons have the same mass independent of the hydrogenic state in which they are excited. The wave function associated with the exciton is

$$\Psi_{ex} = \sum_{k_l k_h} \Psi^{(\kappa,n)}_{k_l k_h} \psi_{k_l}(r_l)\psi_{k_h}(r_h) \tag{12}$$

where ψ_{k_l} and ψ_{k_h} are the Bloch functions associated with the conduction and valence bands and $\Psi_{k_l k_h}^{(\kappa,n)}$ is the Fourier transform of 4.

We now proceed to the more general formulation using essentially the same steps as have been outlined above. Consider the conduction and valence electron extrema given by solutions of the secular determinants

$$\begin{aligned} ||H_{ij}^{e} - E_e\delta_{ij}|| &= 0 \\ ||H_{ij}^{h} - E_h\delta_{ij}|| &= 0, \end{aligned} \tag{13}$$

where the matrix elements H_{ij}^{e} and H_{ij}^{h} are of the general quadratic forms in the momenta:

$$\begin{aligned} H_{ij}^{e} &= \sum_{\alpha\beta} D_{ij}^{e\alpha\beta}(p_e - p_e^{\circ})_\alpha(p_e - p_e^{\circ})_\beta \\ H_{ij}^{h} &= \sum_{\alpha\beta} D_{ij}^{h\alpha\beta}(p_h - p_h^{\circ})_\alpha(p_h - p_h^{\circ})_\beta. \end{aligned} \tag{14}$$

The energy band extrema for conduction and valence bands occur at $\mathbf{p}_e^{0}$ and $\mathbf{p}_h^{0}$ respectively. In an effective mass approximation, the exciton Hamiltonian is

$$\begin{aligned} H_{ii',jj'} &= H_{ij}^{e}\delta_{i'j'} + \delta_{ij}H_{i'j'}^{h} \\ &\quad + V(|\boldsymbol{r}_e - \boldsymbol{r}_h|)\delta_{ij}\delta_{i'j'}, \end{aligned} \tag{15}$$

which in matrix form is written

$$\begin{aligned} \mathcal{H} &= \mathcal{H}^{e}(\boldsymbol{p}_e - \boldsymbol{p}_e^{\circ}) + \mathcal{H}^{h}(\boldsymbol{p}_h - \boldsymbol{p}_h^{\circ}) + \\ &\quad + \mathcal{V}(|\boldsymbol{r}_e - \boldsymbol{r}_h|). \end{aligned} \tag{16}$$

Making the co-ordinate transformation (2), the Hamiltonian (16) becomes

$$\begin{aligned} \mathcal{H} &= \mathcal{H}^{e}(\boldsymbol{p} + \tfrac{1}{2}\boldsymbol{P} - \boldsymbol{p}_e^{\circ}) + \mathcal{H}^{h}(-\boldsymbol{p} + \tfrac{1}{2}\boldsymbol{P} - \boldsymbol{p}_h^{\circ}) + \\ &\quad + \mathcal{V}(r), \end{aligned} \tag{17}$$

where $\mathbf{p}$ and $\mathbf{P}$ are the momenta conjugate to $\mathbf{r}$ and $\boldsymbol{R}$ respectively. The commutativity of $\nabla\mathbf{r}$ with the Hamiltonian (17) once again allows one to look for eigenfunctions of the form

$$\Psi = e^{i\boldsymbol{K}\cdot\boldsymbol{R}}(r), \tag{18}$$

where $\mathbf{K}$ is the exciton wave vector, and $\psi(r)$ satisfies the matrix equation

$$\bar{\mathcal{H}}\psi(r) = E\psi(r), \tag{19}$$

where

$$\begin{aligned} \bar{\mathcal{H}} &= \mathcal{H}^{e}(\boldsymbol{p} + \tfrac{1}{2}\hbar\boldsymbol{K} - \boldsymbol{p}_e^{\circ}) + \\ &\quad + \mathcal{H}^{h}(-\boldsymbol{p} + \tfrac{1}{2}\hbar\boldsymbol{K} - \boldsymbol{p}_h^{\circ}) + \mathcal{V}(r). \end{aligned} \tag{20}$$

In a crystal co-ordinate representation the Wannier equation (19) may be written as

$$\bar{\mathcal{H}}'\psi'(r) = E\psi'(r), \tag{21}$$

where

$$\psi'(r) = \exp\{\tfrac{1}{2}i(h)^{-1}(\boldsymbol{p}_h^{\circ} - \boldsymbol{p}_e^{\circ})\cdot r\}\psi(r), \tag{22}$$

$$\begin{aligned} \bar{\mathcal{H}}' &= \mathcal{H}^{e}(-i\hbar\nabla_r + \tfrac{1}{2}\hbar\,\boldsymbol{\varkappa}) + \\ &\quad + \mathcal{H}^{h}(-i\hbar\nabla_r - \tfrac{1}{2}\hbar\,\boldsymbol{\varkappa}) + \mathcal{V}(r), \end{aligned} \tag{23}$$

and

$$\hbar K = \boldsymbol{p}_e^{\circ} + \boldsymbol{p}_h^{\circ} + \hbar\,\boldsymbol{\varkappa}. \tag{24}$$

Eq. (22) represents a simple displacement of the origin in wave vector space to $(2\hbar)^{-1}$ $(\mathbf{p}_e^{0} - \mathbf{p}_h^{0})$. The matrices $\mathcal{H}^{e}$ and $\mathcal{H}^{h}$ are second order in $\mathbf{p}$ and

VOL. 1 1956/5

may be written in terms of a matrix of effective mass tensors as

$$\mathcal{H}^e(\mathbf{p}) = \tfrac{1}{2}\sum_{\alpha\,\beta}\left(\frac{1}{m_e}\right)_{\alpha\beta} p_\alpha p_\beta \qquad (25)$$

$$\mathcal{H}^h(\mathbf{p}) = \tfrac{1}{2}\sum_{\alpha\,\beta}\left(\frac{1}{m_h}\right)_{\alpha\beta} p_\alpha p_\beta.$$

Substitution of (25) into (23) yields the exciton Hamiltonian in a convenient crystal co-ordinate representation,

$$\mathcal{H}' = \mathcal{H}^\circ - \tfrac{1}{2}i\hbar^2 \sum_{\alpha\,\beta} \varkappa_\alpha \frac{\partial}{\partial r_\beta}\left(\frac{1}{m_e}-\frac{1}{m_h}\right)_{\alpha\beta} +$$

$$+\tfrac{1}{8}\hbar^2 \sum_{\alpha\,\beta} \varkappa_\alpha \varkappa_\beta \left(\frac{1}{m_e}+\frac{1}{m_h}\right)_{\alpha\beta} \qquad (26)$$

where

$$\mathcal{H}^\circ = \mathcal{H}^e(-i\hbar\nabla_r)+\mathcal{H}^h(-i\hbar\nabla_r)+\mathcal{V}(r). \qquad (27)$$

3. EXCITON-BAND STRUCTURE

The exciton spectrum at $\boldsymbol{\varkappa} = 0$ will be given by the solutions of the Wannier equation

$$H^\circ \boldsymbol{\varphi}_n = E_n{}^\circ \boldsymbol{\varphi}_n. \qquad (28)$$

The dispersion relation for the exciton band in the neighborhood will be calculated using the term

$$\mathcal{H}' = -\tfrac{1}{2}i\hbar^2 \sum_{\alpha\,\beta} \varkappa_\alpha \frac{\partial}{\partial r_\beta}\left(\frac{1}{m_e}-\frac{1}{m_h}\right)_{\alpha\beta} \qquad (29)$$

as a perturbation. The eigenfunctions of (28) have either even or odd parity, since the inversion operator commutes with the Hamiltonian independent of whether the crystal is itself invariant under inversion. Hence all first-order terms in the perturbation vanish. The exciton band will accordingly have a critical point at $\boldsymbol{\varkappa} = 0$. The classification of the critical point as maxima, minima, saddle points, etc., will depend in general on the magnitude and sign of the terms in the matrix of effective mass tensors. An exciton of wave vector $\mathbf{K}$ is formed from electron and hole-wave packets with wave vectors $\mathbf{k}_e$ and $\mathbf{k}_h$ respectively such that

$$\boldsymbol{K} = \boldsymbol{k}_e + \boldsymbol{k}_h. \qquad (30)$$

If $\mathbf{K}^0 = \mathbf{k}_e{}^0 + \mathbf{k}_h{}^0$ represents an extremum, an approximate criterion that it will be a minimum is that all $\mathbf{k}_e$ and $\mathbf{k}_h$ satisfying eq. (30) and in the neighborhood $\mathbf{k}_e{}^0$ and $\mathbf{k}_h{}^0$ represent larger energy differences between the electronic energy bands than the difference at $\mathbf{k}_e{}^0$ and $\mathbf{k}_h{}^0$, i.e.

$$|E_e(\boldsymbol{k}_e)-E_h(\boldsymbol{k}_h)| > |E_e(\boldsymbol{k}_e{}^\circ)-E_h(\boldsymbol{k}_h{}^\circ)|, \qquad (31)$$

where $E_e(\mathbf{k})$ and $E_h(\mathbf{k})$ are solutions of the secular determinants (13).

Unfortunately an exact solution of the Wannier eq. (28) can only be obtained for spherical energy surfaces (see eq. 7). However, if the components in the matrices of effective mass tensors $(1/\mathbf{m}_e)$ and $(1/\mathbf{m}_h)$ are known, one may compute the exciton spectrum by a suitable variational calculation. The impurity state problem discussed extensively by KOHN and LUTTINGER[14] may be considered as a special case of the exciton Hamiltonian (27). For example, if one allows the electron mass to become infinite, the acceptor state Hamiltonian results. Similarly the donor-state problem is obtained if one sets the hole mass equal to infinity. The impurity-state spectrum calculated by KOHN and LUTTINGER[14] for donors in Si represents a particular example of the sort of exciton spectrum which one would obtain from nondegenerate ellipsoidal energy surfaces. KOHN and SCHECHTER[15] have applied a variational calculation to the more complex problem arising in the case of degenerate-band impurity-state problems. This Hamiltonian is similar to that for an exciton formed from a degenerate valence band and a spherical-conduction band. Exciton spectra for other combinations of degenerate bands can be obtained by variational procedures similar to those cited above.

A generalization of the *f*-sum rule for the Hamiltonian (27) useful for the application of perturbation theory will now be derived. The commutator of r_γ and H^0 is

$$[r_\gamma, H^\circ] = -\tfrac{1}{2}\sum_{\alpha\,\beta}[p_\alpha p_\beta, r_\gamma]\left(\frac{1}{m_e}+\frac{1}{m_h}\right)_{\alpha\beta}$$

$$= i\hbar \sum_{\alpha} p_\alpha \left(\frac{1}{m_e}+\frac{1}{m_h}\right)_{\alpha\beta}. \qquad (32)$$

Taking the matrix element of this commutator, we get the expression

$$(\varphi_n|r_\gamma|\varphi_{n'}) = -i\hbar \sum_\alpha \frac{\left(\varphi_n|p_\alpha\left(\frac{1}{m_e}+\frac{1}{m_h}\right)_{\alpha\gamma}|\varphi_{n'}\right)}{E_n^\circ - E_{n'}^\circ} \tag{33}$$

which, when substituted into the matrix elements of the commutator

$$[p_\beta, r_\gamma] = -i\hbar\delta_{\beta\gamma}, \tag{34}$$

$$\sum_{n'} [(\varphi_n|p_\beta|\varphi_{n'})(\varphi_{n'}|r_\gamma|\varphi_n) - -(\varphi_n|r_\gamma|\varphi_{n'})(\varphi_{n'}|p_\beta|\varphi_n)] = -i\hbar\delta_{\beta\gamma}, \tag{35}$$

gives the generalized f-sum rule

$$\sum_{\alpha,\, n'}{}' (E_n^\circ - E_{n'}^\circ)^{-1} \Big\{(\varphi_n|p_\beta|\varphi_{n'})\times \times\left(\varphi_{n'}|p_\alpha\left(\frac{1}{m_e}+\frac{1}{m_h}\right)_{\alpha\gamma}|\varphi_n\right)+ +\left(\varphi_n|p_\alpha\left(\frac{1}{m_e}+\frac{1}{m_h}\right)_{\alpha\gamma}|\varphi_{n'}\right)(\varphi_{n'}|p_\beta|\varphi_n)\Big\} = -\delta_{\beta\gamma}. \tag{36}$$

Applying second-order perturbation theory to the Hamiltonian (27) using the term (29) as a perturbation, gives the energy of the nth exciton level as

$$E_n(\boldsymbol{\varkappa}) = E_n^\circ + +\tfrac{1}{8}\hbar^2 \sum_{\alpha\,\beta} \varkappa_\alpha \varkappa_\beta \left(\varphi_n|\left(\frac{1}{m_e}+\frac{1}{m_h}\right)_{\alpha\beta}|\varphi_n\right)+ +\tfrac{1}{4}\hbar^2 \sum_{\substack{\alpha\,\alpha'\\ \beta\,\beta'\\ n'}}{}' \varkappa_\alpha \varkappa_{\alpha'} (E_n^\circ - E_{n'}^\circ)^{-1} \times \times\left(\varphi_n|p_\beta\left(\frac{1}{m_e}-\frac{1}{m_h}\right)_{\alpha\beta}|\varphi_{n'}\right)\times \times\left(\varphi_{n'}|p_{\beta'}\left(\frac{1}{m_e}-\frac{1}{m_h}\right)_{\alpha'\beta'}|\varphi_n\right). \tag{37}$$

For the most general type of degenerate bands, the f-sum rule (36) is of limited help in carrying out the summation (37). However, in a few specific examples, eq. (37) can be summed exactly using (36), and even when the more general case cannot be done exactly, the procedure may be used to get an approximate expression for the exciton dispersion relation. For the general case the dispersion relation for the exciton bands will depend on the quantum state and degeneracy of the exciton.

The exciton eigenfunctions at the band extrema are

$$\Psi_{ex} = \sum_{k_l k_h j j'} \Psi^{(\kappa,n)}_{k_l k_h j j'} \psi_{k_l j}(r_l)\psi_{k_h j'}(r_k) \tag{38}$$

where $\psi_{k_l j}$ and $\psi_{k_h j'}$ are the Bloch functions of the degenerate conduction and valence bands respectively; the subscripts label the wave vectors at the extrema and the degenerate states; and $\Psi^{(\kappa,n)}_{k_l k_h j j'}$ is the Fourier transform of 18. Two subscripts are needed to label the elements of the column matrix $\boldsymbol{\Psi}$, as the Hamiltonian (15) is the direct product of the subspaces for electrons and holes.

In addition to the gross features of the exciton spectrum obtained by solving eq. (28), the levels may show structure. Experimental observations of Gross *et al.*[8] on Cu_2O indicate that at low temperatures the lifting of some degeneracies may indeed occur. The level splittings may arise from magnetic-moment interactions or from higher-order terms in the effective mass approximation. The higher-order terms will tend to mix degenerate minima if these should be present, thus splitting the exciton band degeneracies.

We shall discuss first the exciton fine structure. Provided the spin-orbit interaction is weak, an effective particle moving in a magnetic field has a magnetic moment[18, 19] associated with it of

$$\boldsymbol{\mu} = \frac{K' e\hbar}{2c}\mathcal{J} \tag{39}$$

where $\mathcal{J}$ is the angular momentum matrix associated with the particular degeneracy being considered. K' is a constant which differs from $1/m$ (m is the free electron mass) if there is any orbital angular momentum mixed into the states at the extrema. For the special case of the two-fold degenerate $p^{\frac{1}{2}}$ states in cubic symmetry, $K' \simeq 1/m^*$. The Zeeman splitting of exciton levels given by Samoilovich and Korenblit[20] assumed $K' = 1/m^*$ and hence this treatment needs slight modification to cover the more general case. The fine structure interaction will consist of the dipole-dipole interaction

VOL. 1 1956/5

$$\mathcal{H}_d = \left(\frac{e\hbar}{2c}\right)^2 K_e' K_h' \left\{\frac{r^2 \mathbf{J}_e \cdot \mathbf{J}_h - 3(\mathbf{J}_e \cdot \mathbf{r})(\mathbf{J}_h \cdot \mathbf{r})}{r^5}\right\} \tag{40}$$

as well as the contact term

$$\mathcal{H}_c = -\frac{8\pi}{3}\left(\frac{e\hbar}{2c}\right)^2 K_e' K_h' \mathbf{J}_e \cdot \mathbf{J}_h \delta(r), \tag{41}$$

where the direct products of the vector components of the angular-momentum matrices occur. Under conditions of spherical symmetry, the dipole-dipole interaction vanishes, but the contact term will still give a fine structure splitting to the level. For two spherical energy surfaces the ground states are split into singlet and triplet states, with the singlet state being lower in energy. This problem is completely analogous to that of positronium in free space.

4. DISCUSSION

By far the most serious restriction imposed on the effective mass approximation is the requirement that the interacting potential vary slowly over the lattice spacing. The ground-state binding energy of the exciton, E_1^0 is of the order of $e^2/(2\epsilon\bar{a})$, where $\bar{a}$ is a measure of the dimensions of the exciton. For the potential to be slowly varying we require $\bar{a} >$ lattice constant, or $|E_1^0| < \sim 0{\cdot}2eV$. Hence, in ionic crystals where the observed optical exciton transitions correspond to energies $\sim 0{\cdot}2eV$, one would expect the atomic orbital approximation to be valid. On the other hand, in covalent semiconductors and insulators of large dielectric constant, one expects the effective mass approach outlined above to be the more correct. Even in those cases where the effective mass approximation is applicable, higher-order corrections to the approximation will be effective in splitting some of the remaining degeneracy. For example, if the energy extrema are ellipsoids, the higher-order corrections will mix exciton states formed from the various ellipsoids. This type of correction to the effective mass approximation has been discussed extensively by KOHN and LUTTINGER[14] in connection with the impurity-state problem for donors in Si and Ge. In the general case, the exciton state at the exciton band minima will have zero group velocity, and hence the correction terms will have the crystallographic symmetry of the lattice. The splittings are indicated in the usual group theoretical way by reducing the direct product representation of the wave functions (38) into the representations of the crystallographic point symmetry group. The effect of these correction terms on states different from $\mathbf{K} = 0$ will be that of scattering excitons among the nearly degenerate bands, i.e. $\mathbf{K}$ is no longer a good quantum number. This scattering might be the most effective for producing singlet-triplet-state conversion of excitons formed from spherical electronic-energy surfaces, provided the correction terms are sufficiently large.

It was mentioned in the introduction that optical excitation of the excitonic energy levels is presumably observable experimentally. The selection rule for direct optical transitions is that only transitions to the $\mathbf{K} = 0$ exciton state are possible. This is easily seen since in the filled valence band (the initial state) the wave vectors sum to zero. The absorbed photon contributes a negligible amount of momentum, so that the final exciton state must be at $\mathbf{K} \simeq 0$. Indirect transitions to exciton states at general points in the Brillouin occur by the simultaneous absorption of a photon and emission of a phonon. However, the photon frequency is not unique due to the spread in possible values of the final exciton state, and hence no sharp absorption line spectrum would be observed. The indirect transitions to exciton states would contribute a low energy tail to the absorption curve similar to that discussed by HALL, BARDEEN, and BLATT[21] for indirect transitions to electronic levels. A more careful analysis of absorption coefficients would presumably show an additional low-energy tail due to indirect exciton transitions in cases where indirect electronic transitions are observed.

One source of line width for the optical excitation of excitons is indirect transitions to states $\mathbf{K} \simeq 0$ through simultaneous absorption or induced emission of thermal phonons. The line width expected from this mechanism is $\sim kT$ which corresponds to the temperature dependence and order of magnitude of the line widths observed in the absorption spectra in Cu_2O, CdS, etc.

The simplest band structure conceivable in noncubic substances consists of nondegenerate ellipsoidal extrema. This case may be solved either by the perturbation treatment using the generalized f-sum rule or by a center of mass transformation.

The perturbation calculation will be presented here as a further illustration of the approach. If the symmetry is such that in a given co-ordinate system in wave vector space all effective mass tensors are diagonal (e.g. ⟨100⟩ ellipsoids may all be diagonalized in the same co-ordinate system, however ⟨111⟩ ellipsoids will not all be so diagonalized), then

$$\left(\frac{1}{m_e}\right)_{\alpha\beta} = \frac{1}{m_{e\alpha\beta}}\delta_{\alpha\beta}$$
$$\left(\frac{1}{m_h}\right)_{\alpha\beta} = \frac{1}{m_{h\alpha\beta}}\delta_{\alpha\beta} \qquad (42)$$

and eq. (36) becomes

$$2\left(\frac{1}{m_e}+\frac{1}{m_h}\right)_{\gamma\gamma}\sum_{n'}{}' \frac{(\boldsymbol{\varphi}_n|p_\beta|\boldsymbol{\varphi}_{n'})(\boldsymbol{\varphi}_{n'}|p_\gamma|\boldsymbol{\varphi}_n)}{E_n^\circ - E_{n'}^\circ} = -\delta_{\beta\gamma}. \qquad (43)$$

Eq. (37) may be summed by use of eq. (43) to give

$$E_n(\mathbf{x}) = E_n^\circ + \tfrac{1}{2}\hbar^2 \sum_\alpha \mathbf{x}_\alpha^2 (m_{e\alpha\alpha}+m_{h\alpha\alpha})^{-1}. \qquad (44)$$

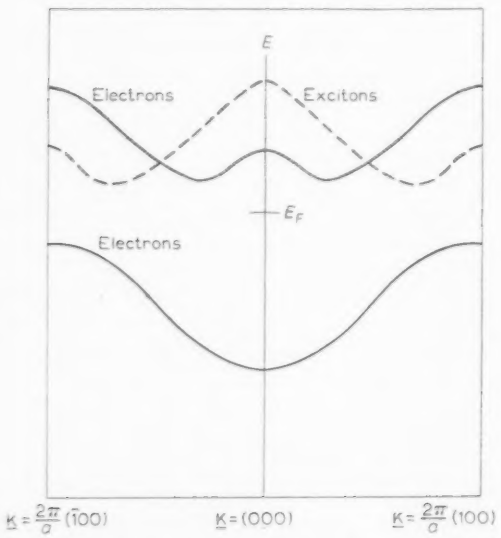


FIG. 1. Schematic drawing of electron- and exciton-band structure in a Brillouin zone with nondegenerate ellipsoidal energy extrema. The exciton-band minima occur at points in the zone which are the sum of the values of the wave vector at the electron and hole extrema. The Fermi energy, E_F, is situated such that the filled lower band represents the ground state of the system. The single exciton band is really an infinite manifold of nearly degenerate levels.

The exciton-band structure in noncubic substances such as CdS may be of the type given by eq. (44). Fig. 1 shows a highly schematic sketch of the exciton-band structure for a case where the electronic extrema are simple ellipsoids at general positions in the Brillouin zone. The single band, labelled excitons, is in reality a manifold of closely spaced bands.

Only order of magnitude calculations have so far been made for the more complicated degeneracies such as those exhibited in the valence-band structure in silicon and germanium. The calculations may, however, often be simplified by the following observation. If one is dealing with a degenerate band valence band and a nondegenerate conduction of light mass, then the non-diagonal terms may be treated as small perturbations. From the known band structure in silicon and germanium[16] order of magnitude estimates have been made for the lowest exciton levels at various minima in the exciton band. A rough estimate of the exciton band structure in Si and Ge is shown in Fig. 2. Estimates of effective masses for excited states[17] were made, based on matrix elements calculated from cyclotron resonance data. These excited state masses have been used in approximating the exciton band structure. The single bands, marked excitons, are once more manifolds of closely spaced levels.

In the interpretation of optical measurements on insulating crystals any observed structure at the absorption-band edge which lies in a region much less than 0·2 eV away from the band edge may be the result of exciton or impurity levels derivable within the effective mass approximation. The distinction between exciton and impurity levels can in principle be made by controlled doping of the crystal. For most of the spectra observed so far, this distinction has not been carefully made. The accurate interpretation of the spectra may yield information about the form and degeneracy of the conduction and valence-electron band structure. In addition, experimental information on the Zeeman splitting may help to identify band degeneracies and to establish the effective magnetic moments of the conduction and valence electrons.

The close analogy between the exciton and positronium suggests that the question of the existence

VOL. 1 1956/5

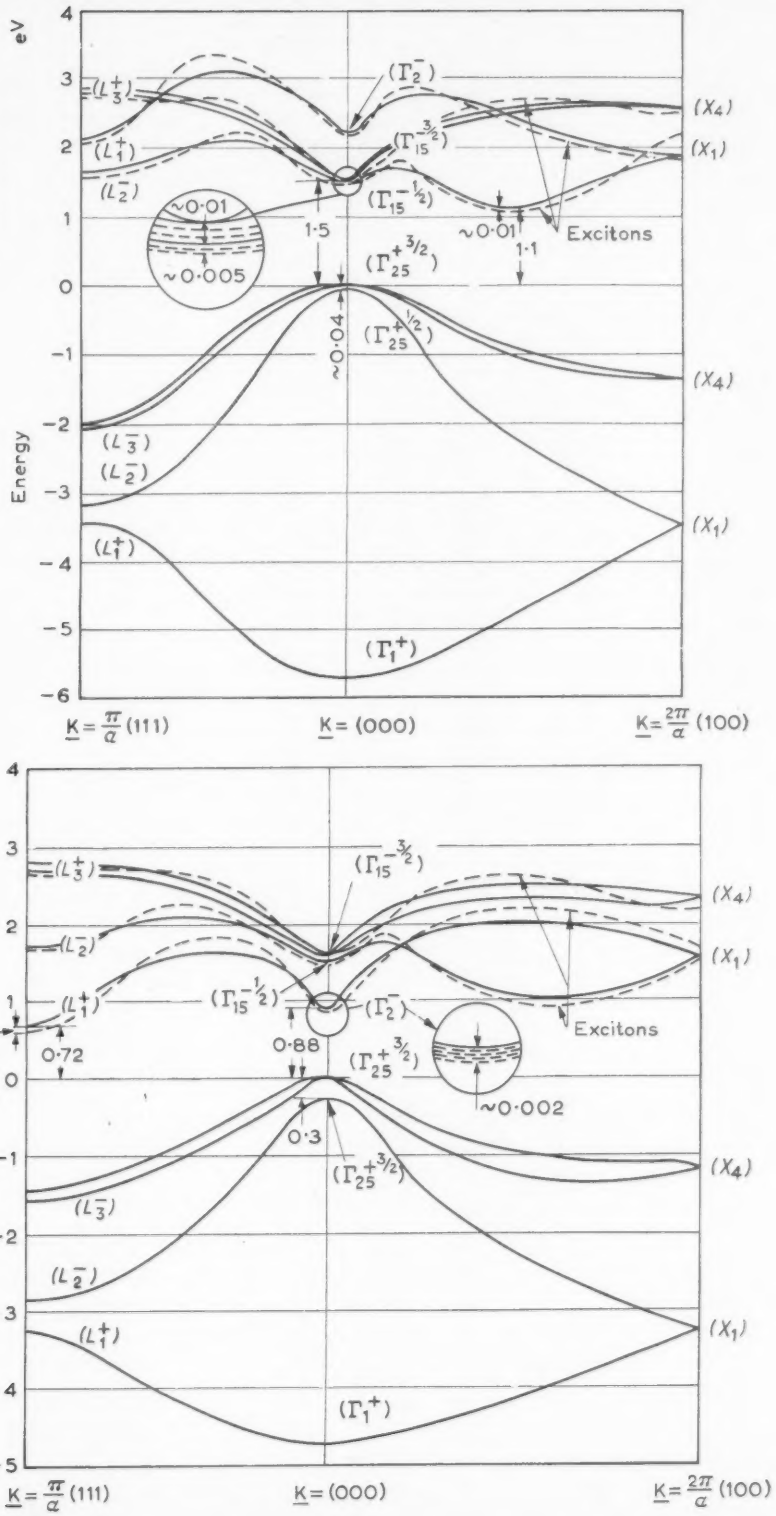


FIG. 2. Schematic diagram of the energy-band structure in (a) silicon and (b) germanium. The binding energies for excitons represent rough estimates based on the parameters known from cyclotron resonance results. The exciton band is an infinite manifold of nearly degenerate bands.

of positronium in insulating solids might be closely related to the considerations which we have discussed. A positron will be weakly bound in the solid, since the largest part of its wave function is confined to the large slowly varying regions of potential between ions. Weak binding implies that the positron effective mass is approximately that of a free electron. The formation of positronium implies that a thermalized positron binds a conduction electron. The only effect of the positron on a filled band is a polarization of the atomic core states. The binding which results from positronium formation in insulating crystals will in most cases be $\sim 0{\cdot}2$ eV. The excitation to a conduction band will in general require more energy than is available from positronium formation. Hence, the existence of positronium in an insulator would thus appear energetically unfavourable. In an amorphous solid where voids are available in which a positron may bind an electron closely, the free-space binding energy of positronium (6·8 eV) is available to extract an electron from a valence band and form positronium. In semiconductors one may have electrons available in the conduction band by thermal excitation or by doping. Here one can always decrease the energy of the system by forming positronium. However, the capture time associated with a positive charge in a semiconductor ($\sim 1\mu$ sec) is much longer than the decay time of positrons in solids. Hence, it seems unlikely that positronium will form even in this event. These speculations are consistent with the experimental observations of BELL and GRAHAM[22] and TALEGDI[23] who could detect the typical free-space positronium decay only in amorphous solids.

Acknowledgements—I would like to thank Professor M. H. COHEN for the many helpful suggestions and criticisms made during the course of this work. I am also grateful to Professor F. REIF for several enlightening discussions.

REFERENCES

1. FRENKEL J. *Phys. Rev.* **37,** 17 (1931).
2. SEITZ F. *Modern Theory of Solids* 414, McGraw-Hill Book Co., New York (1940).
3. OVERHAUSER A. W. *Phys. Rev.* **101,** 1702 (1956).
4. SEITZ F. *Rev. Mod. Phys.* **26,** 7 (1954).
5. WANNIER G. *Phys. Rev.* **52,** 191 (1937).
6. KITTEL C. and MITCHELL A. H. *Phys. Rev.* **96,** 1488 (1954).
7. LUTTINGER J. M. and KOHN W. *Phys. Rev.* **97,** 869 (1955).
8. GROSS E., ZAHARCHENIA B. and REINOV N. *Doklady Akad. Nauk. USSR* **99,** 231 (1954).
9. APFEL J. and HADLEY L. N. *Phys. Rev.* **100,** 1689 (1955).
   NIKITINE S., PERNY G. and SIESKIND M. *Compt. rend.* **238,** 67 (1954).
10. GROSS E. F. and JACOBSON M. A. *Doklady Akad. Nauk. USSR* **102,** 485 (1955).
11. NIKITINE S. *Helv. Phys. Acta* **28,** 307 (1955).
    GROSS E. F. and KAPLIANSKI A. A. *Zhernal Tech. Fiziki* **25,** 2061 (1955).
12. SHOCKLEY W. *Phys. Rev.* **78,** 173 (1950).
13. DRESSELHAUS G. *Phys. Rev.* **100,** 580 (1955).
14. KOHN W. and LUTTINGER J. M. *Phys. Rev.* **98,** 915 (1955).
15. KOHN W. and SCHECHTER D. *Phys. Rev.* **99,** 1903 (1955).
16. DRESSELHAUS G., KIP A. and KITTEL C. *Phys. Rev.* **98,** 368 (1955).
17. DRESSELHAUS G. Thesis, University of California (1955) unpublished.
18. COHEN M. H. (private communication).
19. LUTTINGER J. M. *Phys. Rev.*, **102,** 1030 (1956).
20. SAMOILOVICH and KORENBLIT. *Doklady Akad. Nauk. USSR* **100,** 43 (1955).
21. HALL L., BARDEEN J. and BLATT F. *Phys. Rev* **95,** 559 (1954).
22. BELL R. E. and GRAHAM R. L. *Phys. Rev.* **90,** 644 (1953).
23. TELEGDI V. (to be published).

VOL. 1 1956/5

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 23–26.

RESIDUAL RESISTIVITY vs. SHORT-RANGE ORDER IN Cu_3Au*

A. C. DAMASK

Pitman-Dunn Laboratories, Frankford Arsenal, Philadelphia, Pennsylvania†

(*Received* 18 *May* 1956)

Abstract—The electrical resistivity of Cu_3Au has been measured at 77°K after quenches from various temperatures above T_c, the critical temperature for long-range order. The resistivity decreases with increasing quench temperature up to 485°C, and measurements at 4°K show essentially the same behavior. The origin of this resistance decrease is assigned to short-range order. This decrease of residual resistivity in Cu_3Au with decreasing short-range order is opposite to the results of the author's previous measurements on α brass. Some implications of these two types of resistivity behavior are discussed.

INTRODUCTION

ALTHOUGH it has been assumed by many authors that the electrical resistivity of an ordering alloy decreases with increasing short-range order, little experimental evidence has been presented to justify this assumption. It will be shown that there are two types of resistivity behavior in alloys, one in which the resistivity increases with increasing short-range order, and the other in which the resistivity decreases with increasing short-range order.

A useful abbreviation for the change of resistivity with short-range order will be $d\rho/d\sigma$ where σ is BETHE'S short-range-order coefficient.‡ It should be remembered that σ increases with increasing short-range order, and hence increases with decreasing temperature.

BURNS and QUIMBY[1] measured the resistance at temperature of Cu_3Au above T_c, the critical temperature for long-range order. They found that although $d\rho/dT$, the change of resistance with temperature, had a positive slope at T_c, this slope increased with increasing temperature for about 8°C and then remained constant. They suggested that if this constant resistance slope arose from the thermal part of the resistivity and was subtracted from the observed resistance curve, the difference would be a decrease with increasing temperature, i.e., have a negative slope. If this difference is assigned to changes in short-range order, $d\rho/d\sigma$ would be positive. This contrasts with the author's previous measurements which showed that for α brass $d\rho/d\sigma$ is negative[2].

It was felt that measurements of residual resistivity vs. short-range order should be made on Cu_3Au to see whether $d\rho/d\sigma$ really is opposite in sign to that of α brass, or whether the observations of BURNS and QUIMBY are due to some effect in the thermal part of the resistivity.

SPECIMENS AND PROCEDURE

Proper weight proportions of 99·95 per cent Au and 99·999 per cent Cu were sealed in an evacuated Vycor tube and subjected to the usual melting, shaking, and annealing. The chemical analysis showed that the final alloy was within ¼ atomic per cent of stoichiometric composition. The alloy was swaged and drawn into 0·017-in. wire. The wire was then annealed in vacuo for 24 hours at 750°C. The specimens were prepared by hard-soldering two potential probes of the same alloy onto a 2½-in. wire. Except for one specimen which was used to search for the critical temperature, none was kept at a temperature where it might achieve any degree of long-range order.

A furnace was constructed with a vertical heating

* This work was supported in part by the U.S. Atomic Energy Commission.

† Guest Scientist at Brookhaven National Laboratory, Upton, New York.

‡ σ is chosen rather than α_1 because α_1 is negative and would possibly lead to confusion in the sign of the slope.

chamber ½ in. in diameter and 15 in. long. Separately controllable compensating coils enabled the gradient along the center section to be kept within one-quarter of a degree up to 440°C and one-half of a degree thereafter. The thermocouple which was used to explore the gradient and to measure the temperature was calibrated at the freezing points of National Bureau of Standards lead and zinc.

The specimen wire was vacuum-sealed in a pyrex tube with a small copper rod to add mass, annealed in the furnace, and then was quenched by dropping the tube directly from the furnace into a container of ice water which had a metal plate at the bottom to assist in fracture of the tube. It was found that the tube shattered better when it was scored heavily with a file. Only quenches which completely shattered the glass were accepted. When a quench was acceptable the wire was connected into the circuit and its resistance measured in liquid nitrogen by comparing the potential drop to that of a dummy wire.

Since the resistivity changes that were observed were quite small, it was found that the cold-work resistivity of a small bend in the wire, caused by steam pressure or breaking glass, rendered the result and the wire useless. Because of this uncertainty of whether or not the wire had been slightly cold worked during a quench, the measurements were made by a procedure such as the following: a wire was quenched from 395°C and its resistance was measured. The wire was then quenched from 400°C and measured. It was then quenched again from 395°C and its resistance was checked with the first 395°C quench. This process of repeatedly doubling back to previously validated results enabled deviations owing to cold work to be spotted immediately. If cold work occurred the wire was discarded. It was found that results in the 395–435°C range were readily reproducible and yielded a smooth slope. The measurement of resistivity changes in this range proved to be more accurate than the dimensional measurements of the sample. Therefore, when starting with a new wire, quenches were first made in this range, and the resulting resistivity slope was checked to see if a linear displacement in resistance would superimpose the new slope on the slope obtained with the previous wires.

VOL. 1 1956/5

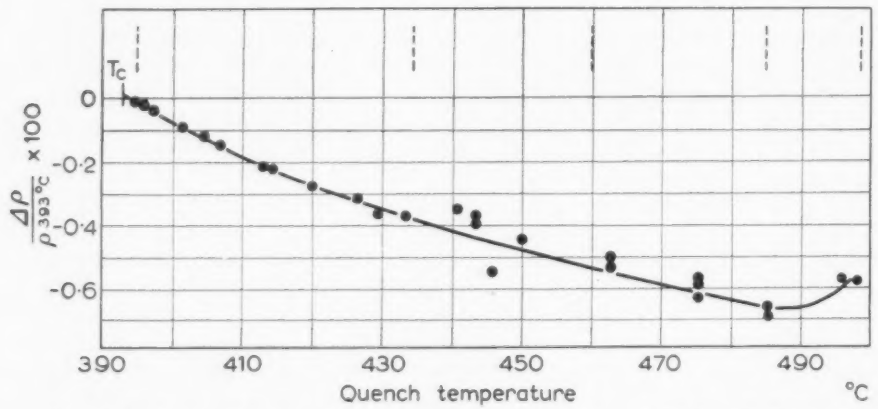


FIG. 1. Resistance changes at 77°K vs. quench temperature of Cu_3Au. Ordinate scale is in tenths of one per cent below the value at the critical temperature.

RESULTS

The results are shown in Fig. 1. They can roughly be divided into four separate parts indicated by the dashed lines. The low temperature range, 395–435°C, gave readily reproducible results, and it was to this slope that the resistivity of a new wire was normalized.

The resistivity of a separate wire was normalized to this slope and was then quenched from 394 and 392°C respectively. It was found that the resistivity after the 394°C quench was in line with the slope shown, but the resistivity after the 392°C quench had fallen by several per cent. The critical temperature was chosen between these two temperatures at 393°C. The resistivity curve in Fig. 1 is extrapolated to 393°C, and the value at that temperature was used as the zero-point resistivity for the calculation of the percentage change shown on the ordinate.

In the preliminary experiments a quench from 395°C was repeated with times in the furnace of 1/2, 1, and 3 hours. It was found that there was no change of resistivity with these increasing furnace times. Since equilibrium was achieved by a 1/2-hour annealing time at the lowest temperature, this time was used throughout.

Resistance measurements were made at 4°K on a specimen quenched respectively from the temperatures 403, 416, and 432°C, after checking its resistance at liquid nitrogen. The slope of the curve at 4°K was about 2 per cent steeper than the slope at liquid-nitrogen temperature, which showed that there is not much effect of short-range order on the thermal part of the resistivity up to 77°K. The resistivity at 4°K of the specimen quenched from 403°C was $9{\cdot}6\times10^{-6}$ ohm-cm. The resistivity of this same specimen at liquid nitrogen was $10{\cdot}7\times10^{-6}$ ohm-cm. This is a slightly larger percentage increase than that reported by PASSAGLIA and LOVE[3], on a specimen quenched from 600°C.

The range 435–460°C was characterized by a great deal of scatter as shown by sample data in Fig. 1, while above 460° fairly good reproducibility was again obtained. Repeated attempts were made to get uniform data in the 435–460°C range by approaching from both high and low temperatures, but the scatter persisted. A suggested reason for this is the following: The lower the temperature the longer is the relaxation time for atomic migration, and the higher the temperature the more readily the glass shatters. The scatter in this range may well be the result of these two competing processes. However, in spite of the scatter it is evident that the resistivity is still decreasing with increasing temperature of quench.

While the reproducibility above 460°C is not quite as good as that of the low-temperature range, it does clearly show the direction of the curve. At 496–8°C the data are higher than those of the preceding temperature. The multiple points above 460°C arise from retracings of that portion of the curve, which verified that the increase at 496°C was real. This increase could not be pursued to higher temperatures because the drastic quenches above this temperature damaged the wires more frequently. Heavier wires were not used for fear of not quenching in all of the effects. The increase at 496°C does show, however, that the limit of the quench had not been reached at lower temperatures.

GIBSON[4] has shown that the resistivity of Cu_3Au just above T_c is about 20 per cent higher than alloy resistivity. Therefore the resistivity due to short-range order would not be expected to alter its direction from the steady decrease to infinite temperature. The increase at 496°C is therefore probably due to some effect other than short-range order. If this is so, a slight correction of the short-range order curve of Fig. 1 may be necessary. This would result in a slightly greater decrease due to short-range order than that shown.

DISCUSSION

It has now been shown experimentally that resistivity of Cu_3Au increases with increasing short-range order while that of α brass decreases with increasing short-range order.

Scientific literature contains many peculiar observations of resistivity changes arising from damage to alloys. The demonstration of the existence of two types of resistivity behavior in alloys characterized by opposite signs of $d\rho/d\sigma$ prompts speculation on alternative interpretations of some earlier experiments and possible future observations.

When α brass is irradiated at room temperature with neutrons, the resistivity decreases. Experiments performed at Brookhaven have shown that this decrease arises from increased short-range order[2, 5]. It is therefore tempting to suspect short-range ordering in other alloys which decrease in resistivity under similar conditions. However, an alloy with a positive $d\rho/d\sigma$ would increase in resistivity if the neutron-produced defects increased short-range order. An increase in resistivity occurring in an irradiated alloy therefore should not be assigned entirely to damage, but should be examined for possible evidence of short-range order.

Cold work will decrease short-range order. For the case of α brass with a negative value of $d\rho/d\sigma$, the destruction of short-range order will increase the resistivity. It is known that the cold-work resistivity in α brass anneals in two distinct decreasing steps, and it has been shown that the low-temperature step is due to the re-establishment of equilibrium short-range order[2]. This same phenomenon occurs in Ag-Zn and has a similar interpretation[6]. In an alloy with a positive $d\rho/d\sigma$ the destruction of short-range order should cause a decrease in resistivity, although it may possibly be masked by an increase due to dislocations. Annealing should show two steps as in α brass but in this case the low temperature step should be an increase in resistivity as equilibrium short-range order is achieved. This behavior would be

observable only in an alloy with T_c low enough so that the annealing would not cause any long-range order. (This latter requirement was demonstrated by isochronally annealing cold-worked Cu_3Au. Only decreasing resistivity was observed. This indicated that in spite of the cold-work damage enough long-range ordering took place to mask any increase in resistivity due to short-range order.)

BLEWITT and KOEHLER[7] measured the resistivity of ordered Cu_3Au as a function of strain. For strains up to 0·2 they observed a small decrease in resistivity. They interpreted this as a decrease in the amount of antiphase-domain wall, but an alternative explanation may now be given. Since dislocations move in pairs through an ordered domain there should be a decrease of short-range order at the boundaries, i.e., an *increase* in the amount of domain wall, prior to a substantial decrease of long-range order within the domains[8, 9, 10]. This early decrease of short-range order could account for the observed resistivity decrease. This might be verified by repeating their experiments on a specimen quenched from 400°C.

GIBSON'S[4] analysis indicates that Au-Ag may also have a positive $d\rho/d\sigma$. If this were true, vacancy enhanced ordering might be a factor, in addition to COTTRELL'S explanation[10], in interpreting the observations of AARTS and JARVIS[11], which showed that an Au-Ag alloy, deformed at 80°K, increased in resistivity after annealing at room temperature.

The author does not intend to suggest that all resistivity anomalies in alloys will be explained by this mechanism, but rather that the existence of alloys with opposite signs of $d\rho/d\sigma$ indicates the desirability of knowing the sign of $d\rho/d\sigma$ in future experiments on the resistivity properties of alloys.

Acknowledgements—Appreciation is expressed to RONALD LARSEN for his valuable assistance in the construction of the furnace and in the taking of preliminary data. Appreciation is also expressed to G. J. DIENES for helpful discussions.

REFERENCES

1. BURNS F. P. and QUIMBY S. L. *Phys. Rev.* **97,** 1567 (1955).
2. DAMASK A. C. *J. Appl. Phys.* **27,** 610 (1956).
3. PASSAGLIA E. and LOVE W. P. *Phys. Rev.* **98,** 1006 (1955).
4. GIBSON J. A. *J. Phys. Chem. Solids* **1,** 27 (1956).
5. ROSENBLATT D. B., SMOLUCHOWSKI R. and DIENES G. J. *J. Appl. Phys.* **26,** 1044 (1955).
6. DAMASK A. C. *Acta Met.* **4,** 215 (1956).
7. BLEWITT T. H. and KOEHLER J. S. *Plastic Deformation of Crystalline Solids* (Symposium Pittsburgh, 1950).
8. MUTO T. and TAGAKI Y. *Solid State Physics, V.* I, 277 (Academic Press, New York, 1955).
9. FISHER J. C. *Acta Met.* **2,** 9 (1954).
10. COTTRELL A. H. *Relation of Properties to Microstructure* (ASM Cleveland, 1954).
11. AARTS W. H. and JARVIS R. K. *Acta Met.* **2,** 87 (1954).

VOL.
1
1956/5

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 27–34.

THE EFFECT OF SHORT-RANGE ORDER ON RESIDUAL RESISTIVITY*

JOHN B. GIBSON

Brookhaven National Laboratory, Upton, New York

(*Received* 18 *May* 1956)

Abstract—A general theory, based on the Nordheim approximation, is given for the residual resistivity of binary solid solutions with short-range order but no long-range order. For Cu_3Au, this theory predicts about an 18 per cent increase in resistivity with short-range order just above the ordering temperature, in agreement with experiment. For α brass, the agreement is poorer, an 0·4 per cent increase in resistivity being predicted in comparison with a 12 per cent decrease observed. Qualitative considerations are given for estimating the effect of short-range order for an arbitrary solid solution. In general, the short-range order may be expected to decrease the resistivity of a binary solid solution whose constituents have different valence when the Fermi surface passes close to a Brillouin zone boundary for the random lattice. An increase should occur when the Fermi surface is close to a zone boundary for the superlattice.

INTRODUCTION

METALLIC conduction arises from the effect of an electric field on conduction electrons. These are accelerated until scattered by some deviation from periodicity in the potential. Deviations from periodicity are caused by vibrations, defects, impurities, etc. Usually the contributions to resistance of different scattering mechanisms are additive (MATTHIESSEN'S Rule). Calculations of residual resistivity (temperature independent) have followed the general method of NORDHEIM,[1] who computed the residual resistivity of a disordered alloy of atoms of nearly the same atomic number. MOTT[2] relaxed the requirement on atomic number. Various workers[3] have taken into account long-range order, MURAKAMI[4] has considered short-range order, and KRIVOGLAZ and MATYSINA[5] have discussed the effects of both long-range and short-range order on resistivity. The latter authors also show that the collective electron picture introduces no essential change from the Nordheim method in the calculation of residual resistivity.

MURAKAMI evaluated the effect of short-range order on residual resistivity for the case of a simple cubic solid solution, and computed both the short-range order vs. temperature and its effect on resistivity. The effect of short-range order on the residual resistivity of both $AuCu_3$ and α brass has recently been measured by DAMASK.[6] The calculation of MURAKAMI for the simple cubic case does not apply to $AuCu_3$ or α brass, so these cases are considered. The general method of MURAKAMI was improved. In this work the short-range order was assumed to be known and the effect of this order on residual resistivity calculated. Effects of crystal structure and scattering mechanism are considered. Qualitative considerations are given that may enable one to estimate easily the effect of short-range order on the resistivity of an arbitrary material.

EXTENSION OF THE NORDHEIM THEORY TO SHORT-RANGE ORDER

The basis of the Nordheim Theory is the expansion of electronic wave functions in terms of solutions for some periodic potential $V_0(\mathbf{r})$. By HOUSTON'S theorem[7] the effect of an electric field on these states is to translate their occupancy in the $\mathbf{k}$ space uniformly in the direction of the field. The difference between the actual potential

* Work done under contract with the U.S. Atomic Energy Commission.

$V(\boldsymbol{r})$ and the periodic potential $V_0(\boldsymbol{r})$, used to determine the zero-order states, scatters electrons in these states. The matrix of the perturbing potential is given by

$$V'_{\boldsymbol{kk'}} = \int \psi_{\boldsymbol{k'}}^* [V(\boldsymbol{r}) - V_0(\boldsymbol{r})] \psi_{\boldsymbol{k}} \, d\tau, \tag{1}$$

where $\psi_{\boldsymbol{k}}$ is the initial wave function, and $\psi_{\boldsymbol{k'}}$ is the final wave function, both for the periodic potential $V_0(\boldsymbol{r})$.

Suppose a binary solid solution of A- and B-type atoms has a potential V_A in cells containing A atoms and V_B in cells containing B atoms. Then, following NORDHEIM, we define

$$V_0(\boldsymbol{r}) = x_A V_A + x_B V_B, \tag{2}$$

where the x's are the atom fractions of the constituents. This choice makes the diagonal elements of the matrix (1) zero. The square of the matrix element can be written as

$$|V'_{\boldsymbol{kk'}}|^2 = N^{-2} \sum_{\boldsymbol{nn'}} e^{i(\boldsymbol{k}-\boldsymbol{k'})\cdot(\boldsymbol{n}-\boldsymbol{n'})} (V_{\boldsymbol{n}}^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})(V_{\boldsymbol{n'}}^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})^*, \tag{3}$$

where

$$V_{\boldsymbol{n}}^{\boldsymbol{kk'}} = \int u_{\boldsymbol{k'}} u_{\boldsymbol{k}} V_{\boldsymbol{n}} e^{i(\boldsymbol{k}-\boldsymbol{k'})\cdot\boldsymbol{r}} \, d\tau_0. \tag{4}$$

$V_{\boldsymbol{n}}$ is the potential of the type of atom in the $\boldsymbol{n}$th cell. Bloch states are written in the form $\psi_{\boldsymbol{k}} = N^{-\frac{1}{2}} u_{\boldsymbol{k}}(\boldsymbol{r}) \, e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ where $u_{\boldsymbol{k}}(\boldsymbol{r})$ is periodic, and N is the number of atoms in the crystal. The vectors $\boldsymbol{n}$ and $\boldsymbol{n'}$ go over the lattice sites.

If the sum in eq. (3) is carried out for constant $\boldsymbol{n}-\boldsymbol{n'}$ first, one obtains

$$|V'_{\boldsymbol{kk'}}|^2 = N^{-1} \sum_{\boldsymbol{n''}} e^{i(\boldsymbol{k}-\boldsymbol{k'})\cdot\boldsymbol{n''}} \overline{(V_{\boldsymbol{n}}^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})(V_{\boldsymbol{n'}}^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk}})^*}. \tag{5}$$

The bar designates the average over all pairs of atoms $\boldsymbol{n}$ and $\boldsymbol{n'}$ for which $\boldsymbol{n}-\boldsymbol{n'} = \boldsymbol{n''}$. For the case of no short-range order NORDHEIM showed that only the pairs for which $\boldsymbol{n}-\boldsymbol{n'} = 0$ have nonzero averages. MURAKAMI carried out the averages for the case of short-range order. We will follow his method, using the notation of COWLEY[6] for the definition of short-range-order parameters. In preparation for calculating the average in eq. (5), the short-range-order parameters will now be introduced.

If the origin is occupied by a B atom and p_{lmn} is the probability that a site with co-ordinates $l\ m\ n$ is occupied by an A atom, the short-range-order parameter α_{lmn} is defined as

$$\alpha_{lmn} = 1 - \frac{p_{lmn}}{x_A}. \tag{6}$$

The probability that the site with co-ordinates $l\ m\ n$ is occupied by a B atom is given by

$$\begin{aligned} &x_B(1-\alpha_{lmn}), \text{ where an } A \text{ atom is at the origin, and} \\ &x_B + x_A\alpha_{lmn}, \text{ where a } B \text{ atom is at the origin.} \end{aligned} \tag{7}$$

The next step is to evaluate

$$\overline{(V_{\boldsymbol{n}}^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})(V_{\boldsymbol{n'}}^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})^*}, \tag{8}$$

setting $\boldsymbol{n}-\boldsymbol{n'}$ equal to l,m,n. At site $\boldsymbol{n}$, the atom will either be an A type or B type, similarly for site $\boldsymbol{n'}$. Thus there are just four cases for a given $\boldsymbol{n}-\boldsymbol{n'}$ and the average (8) will be the sum of the four functions weighted by the probability of their occurrence. Thus

$$\begin{aligned} &\overline{(V_{\boldsymbol{n}}^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})(V_{\boldsymbol{n'}}^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})^*} = \\ &x_A(x_B - x_B\alpha_{lmn})(V_B^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})(V_A^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})^* + \\ &x_B(x_B + x_A\alpha_{lmn})(V_B^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})(V_B^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})^* + \\ &x_A(1 - x_B + x_B\alpha_{lmn})(V_A^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})(V_A^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})^* + \\ &x_B(1 - x_B - x_A\alpha_{lmn})(V_A^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})(V_B^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})^* + \\ &= |x_A(V_A^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}}) + x_B(V_B^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})|^2 + \\ &\quad + x_A x_B \alpha_{lmn} |V_A^{\boldsymbol{kk'}} - V_B^{\boldsymbol{kk'}}|^2. \end{aligned} \tag{9}$$

From eq. 2

$$x_A(V_A^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}}) + x_B(V_B^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}}) = 0, \tag{10}$$

so

$$\begin{aligned} &\overline{(V_{\boldsymbol{n}}^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})(V_{\boldsymbol{n'}}^{\boldsymbol{kk'}} - V_0^{\boldsymbol{kk'}})^*} \\ &\quad = x_A x_B \alpha_{lmn} |V_A^{\boldsymbol{kk'}} - V_B^{\boldsymbol{kk'}}|^2. \end{aligned} \tag{11}$$

By symmetry α_{lmn} is the same for all atoms in a shell. (Sites of a shell are obtained by performing all members of the point group of the lattice on one site.) Indexing shells by i one obtains

$$|V'_{\boldsymbol{kk'}}|^2 = N^{-1} x_A x_B |V_A^{\boldsymbol{kk'}} - V_B^{\boldsymbol{kk'}}|^2 \sum_{i,\boldsymbol{n}_i} \alpha_i e^{i(\boldsymbol{k}-\boldsymbol{k'})\cdot\boldsymbol{n}_i}, \tag{12}$$

where the $\boldsymbol{n}_i$ give the atoms in the ith shell.

In the cases under consideration $|V_A^{\boldsymbol{kk'}} - V_B^{\boldsymbol{kk'}}|^2$ will be assumed to depend only on the magnitude of $\boldsymbol{K} = \boldsymbol{k}-\boldsymbol{k'}$, spherical energy surfaces also being assumed. Because of the exponential $e^{i\boldsymbol{k}\cdot\boldsymbol{n}_i}$ in eq. (12) $|V'_{\boldsymbol{kk'}}|^2$ is not dependent on $|\boldsymbol{K}|$ alone. The requirement of azimuthally independent scattering is necessary to

VOL. 1 1956/5

use the usual simple method of computing resistivity as was done by MURAKAMI[4]. The usual formula can be generalized to include the present case of azimuthally dependent scattering by using a variational principle of KOHLER[9].

As discussed in Appendix A for cubic metals with spherical energy bands, KOHLER's principle predicts that the resistivity is proportional to the integral

$$\int\int |V_{kk'}|^2(1-\cos\theta)\, ds_k ds_{k'}, \tag{13}$$

where ν is the angle between k and k'. Both integrations are over the Fermi surface. This is distinct from the usual formula for resistivity where integration is performed only over the final state Fermi surface.

The region of integration in (13) is invariant to rotations of $K = k-k'$, so only that part of the integrand invariant to this operation contributes to the integral[10]. Inserting (12) into (13) and recalling our assumption that $|V_A^{kk'}-V_B^{kk'}|$ depends only on $|K|$, one sees that solely that part of $e^{ik\cdot n_i}$ invariant to rotations of K contributes to the integral. This invariant part is $j_0(Kn_i)$. $j_l(x)$ is the spherical Bessel function of l^{th} order[11].

Thus, the expression for resistivity becomes

$$\rho \propto x_A x_B \int\int |V_A^{kk'}-V_B^{kk'}|^2(1-\cos\theta) \sum_i \alpha_i c_i j_0(kn_i)\, ds_k ds_{k'}, \tag{14}$$

where c_i is the number of atoms in the ith shell.

One may reduce (14) to an integration over K. Since $dS_{k'} = 2\pi \sin\theta k'^2 d\theta$, $k'^2 = k^2$, and $K^2 = k^2+k'^2-2kk'\cos\theta = 2k^2(1-\cos\theta)$, one obtains

$$(1-\cos\theta)dS_{k'} = \pi K^3 dK/k^2. \tag{15}$$

The integrand is now independent of k so that one integration in (14) can be performed at once (i.e. $\int dS_k = 4\pi k^2$). The range of K is from 0 to $2k$. Expression (14) becomes

$$\rho \propto x_A x_B 4\pi^2 \int_0^{2k} |V_A^{kk'}-V_B^{kk'}|^2 \sum_i \alpha_i c_i j_0(Kn_i) K^3\, dK. \tag{16}$$

This expression is similar to the formula for the total cross-section for X-ray scattering for the case of short-range order in a powder specimen,[12] but contains an extra term $1-\cos\theta$ which arises from the fact that large angle collisions are most effective in limiting the current than are small angle ones. Eq. (16) is, in fact, proportional to the transport cross-section.

QUALITATIVE DISCUSSION OF THE EFFECT OF SHORT-RANGE ORDER ON RESISTIVITY

It is now possible to separate the effect of short-range order on resistivity into two parts, one primarily dependent on the crystal structure and ordering scheme, the other on the properties of the constituent atoms. Let us first discuss the function $F(K) \equiv \sum_i \alpha_i e^{ik\cdot n_i}$ and its spherical average $\overline{F(K)} \equiv \sum_i \alpha_i c_i j_0(Kn_i)$. In the absence of short-range order $\alpha_i = 0 (i > 0)$, α_0 always equals one, and therefore $F(K) = 1$. As short-range order develops, $F(K)$ increases for $K/2\pi$ near the superlattice reflections and decreases for $K/2\pi$ near the fundamental reflections. For perfect long-range order $F(K)$ consists of a sum of δ functions located at superlattice reflections. Of interest is the ratio of the resistivity of a material with short-range order to that of the disordered material. Thus integral (16) is performed twice, once with $\overline{F(K)} = 1$ and once with $\overline{F(K)}$ appropriate to whatever short-range order exists. In Fig. 1 the solid curve shows $K(\overline{FK})$ vs. $Ka/2\pi$ calculated for α's measured by COWLEY for AuCu$_3$ at $T = 405°$C, about 10 degrees above the critical temperature. Peaks occur in the solid curve at $Ka/2\pi = 1{\cdot}0$ and $1{\cdot}414$, the first two superlattice reflections, while a dip occurs at $1{\cdot}732$, the first fundamental reflection. The dotted line is the same function for no short-range order.

In order to determine qualitatively whether short-range order increases or decreases residual resistivity one follows this procedure. Locate the peak, if any, in the part of the integrand of (16) $K^2|V_A^{kk'}-V_B^{kk'}|^2$. Now notice whether $KF(K)$ or K is larger in this region. If there is a peak in $K^2|V_A^{kk'}-V_B^{kk'}|^2$ and $KF(K)$ is appreciably different from K, a prediction of the effect of short-range order on residual resistivity can be made. If $KF(K)$ is larger than K in the region of the peak the effect of short-range order will be to increase residual resistivity from that of the disordered material. If $KF(K)$ is smaller than K in the region of the peak of $K^2|V_A^{kk'}-V_B^{kk'}|^2$ the effect is the opposite.

Even if the short-range order of a material has never been measured, perhaps some information about the effect considered here can be found. If the peak in $K^2|V_A^{kk'}-V_B^{kk'}|^2$ is in the region of a superlattice reflection $KF(K)$ will be greater than

K, and if the peak is in the region of a fundamental reflection $KF(K)$ will be smaller than K. It would be a mistake to attach too much weight to estimates made in this manner, since one part of the range of integration of (16) may influence in one direction, and another part of the range in the opposite way.

RESULTS

A. Mott's Approximation Applied to $AuCu_3$

The function $|V_A^{kk'}-V_B^{kk'}|^2$ also arises in the calculation of the residual resistivity of a disordered alloy. MOTT[2] extended the original work of NORDHEIM,[1] and treated the case of a solid solution of monovalent constituents. Recent measurements have been made of the short-range order[8] of $AuCu_3$, and of the residual resistivity[6] as a function of quench temperatures above the critical temperature. Both Au and Cu are monovalent, therefore MOTT's approximation for $|V_A^{kk'}-V_B^{kk'}|^2$ may be used to calculate the residual resistance of $AuCu_3$.

MOTT considered the case of a solid solution of monovalent atoms. He neglected the p part of the wave functions, assuming $\Psi^A_{\boldsymbol{k}}=N^{-\frac{1}{2}}\,u_A^{(r)}\,e^{i\boldsymbol{k}.\boldsymbol{r}}$, where $u_A(r)$, the periodic part of the wave function, is independent of $\boldsymbol{k}$, similarly he also assumed that $u_A(r_S^A)=u_B(r_S^B)$ where r_S^A and r_S^B are the radii of atomic spheres of an A and B type atom, respectively. Under these assumptions the problem of finding the wave functions in a solid solution consists of matching the exponents on the boundary of the cells. This problem is equivalent to one with potential constant in each cell but differing in different cell types. The magnitude of the potential in a cell is adjusted to give the proper k for electrons with the Fermi energy of the pure constituent of that cell.

Assuming further that $r_S^A=r_S^B=r_S$ one finds the value of $V_A^{kk'}-V_B^{kk'}$ as follows:

$$V_A^{kk'}-V_B^{kk'}=(4\pi r_s^3/3)^{-1}(E_F^B-E_F^A)\int_0^{2\pi}\int_0^{\pi}\int_0^{r_s}e^{i\boldsymbol{k}.\boldsymbol{r}}\sin\theta\,d\,\theta\,d\varphi\,r^2\,dr$$
$$=3(E_F^B-E_F^A)j_1(Kr_s)/Kr_s. \qquad (17)$$

The Fermi energy of pure A is E_F^A. The function $|V_A^{kk'}-V_B^{kk'}|^2\ K^2a^2/(2\pi)^2$ for this model is plotted in Fig. 2, (labelled "flat bottomed") for a face-centered cubic solid solution.

Using eq. (17) one can evaluate (16) for the resistivity for the case of monovalent solid solutions as follows:

$$\rho\propto\sum_i\alpha_i c_i I(z_i,x), \qquad (18)$$

where $x=2kr_s$, $z_i=n_i/r$, and

$$\begin{aligned}I(z,x)&\equiv z^{-1}\int_0^x\sin zy(\sin y-y\cos y)^2y^{-4}\,dy\\&=-1/6-(1/12x^2)\cos 2x+\\&+[(z^2/12)-(1/2)-(1/6x^2)](\sin zx)/zx+\\&+[(1/6)-(z^2/12)](\cos 2x\sin zx)/zx+\\&+\left(\cos zx+2\frac{\sin zx}{zx}\right)\left(\frac{\cos 2x}{(2x)^2}+\frac{\sin 2x}{2x}\right)/3\\&+\left(\frac{1}{2}-\frac{z^2}{12}\right)G(zx)+\frac{z^3-6z-4}{24z}G(|2+z|x)+\\&+\frac{z^3-6z+4}{24z}G(|z-2|x),\end{aligned} \qquad (19)$$

where $G(x)=Ci(x)-\ln\gamma x=-\int_0^x\frac{1-\cos t}{t}\,dt$

$Ci(x)$ is the cosine integral and $\gamma=1{\cdot}781\ldots$. For monovalent face-centered-cubic solid solutions (e.g. $AuCu_3$) $x=3{\cdot}8382$ and

$$z_i=(1{\cdot}2794)\{0,\sqrt{2},\sqrt{4},\sqrt{6}\ldots\}. \qquad (20)$$

Table 1 gives $I(z_i,x)$ for these parameters as well as $\alpha_i c_i$ measured by COWLEY.[8] The ratio of the

Table 1

i	Co-ords.	$c_i\alpha_i$ for 405°C	$c_i\alpha_i$ for 460°C	$c_i\alpha_i$ for 550°C	$I(z_i,x)$	$j_0^2-j_1^2$
0	0,0,0	+1·000	+1·000	+1·00	+0·8611	+1·0000
1	1,1,0	−1·82	−1·78	−1·57	−0·0465	−0·0029
2	2,0,0	+1·12	+1·03	+0·63	+0·0625	+0·0026
3	2,1,1	+0·22	+0·46	+0·62	−0·0023	−0·0178
4	2,2,0	+1·14	+0·82	+0·54	−0·0035	+0·0158
5	3,1,0	−1·27	−1·18	−0·77	−0·0002	+0·0028
6	2,2,2	+0·20	+0·06	−0·07	+0·0005	−0·0114
7	3,2,1	−0·77	−0·38	−0·14	−0·0009	−0·0019
8	4,0,0	+0·29	+0·25	+0·14	−0·0014	+0·0077
9a	3,3,0	−0·31	−0·26	−0·13	−0·0005	+0·0050
9b	4,1,1	+0·26	+0·48	+0·46	−0·0005	+0·0050
10	4,2,0	+0·62	+0·60	+0·17	+0·0004	−0·0032

The values of α_i, short-range-order parameters, are those measured by COWLEY[8] on $AuCu_3$ at temperatures given. The number of atoms in the ith shell is c_i. The function $I(z_i,x)$ is used in the calculation of the residual resistivity of monovalent solid solutions, while similarly $j_0^2(Kn_i)-j_1^2(Kn_i)$ is used for solid solutions with constituents of different valence.

VOL. 1 1956/57

residual resistivity for a given state of order to the residual resistivity of a disordered alloy is

$$\rho/\rho_0 = \sum_i \alpha_i c_i I(z_i,x)/I(0,x). \qquad (21)$$

From eq. (6), $c_0\alpha_0=1$, independent of short-range order. For quench temperatures of 405° C, 460° C and 550° C (using COWLEY's values of α_i), the ratio ρ/ρ_0 is found to be 1·177, 1·167, and 1·127 respectively. Refer to Figs. 1 and 2. The reason the ratio ρ/ρ_0 is greater than one is that the large values of $K\sum_i \alpha_i c_i j_0(Kn_i)$ and $|V_A^{kk'}-V_B^{kk'}|^2 K^2a^2/(2\pi)^2$ occur in the same region, about the first super lattice distance $Ka/2\pi = 1$. Thus the integral (14) is greater for short-range order than for no short-range order.

DAMASK[6] measures a decrease of 0·4 per cent. in residual resistivity of $AuCu_3$ as the quench temperature is raised from 405° C to 460° C. The result of this calculation, a 1 per cent. decrease between 405° C and 460 C°, is in good agreement with this experimental result.

B. Nordheim's Approximation Applied to α brass

NORDHEIM's method is suitable for solid solutions of nearly the same atomic number, for example α brass. The average potential is taken to be the average of the atomic potentials weighted by the atom fractions. The zero-order wave functions used in the calculation are Bloch waves for this potential. Scattering is concentrated near the nuclei so that $V_B^{kk'}-V_B^{kk'}$ is nearly independent of K. This assumption of $V_A^{kk'}-V_B^{kk'}$ = constant (isotropic scattering) has been used by SEITZ[13] and MURAKAMI.[4] For this case the function $|V_A^{kk'}-V_B^{kk'}|^2K^2a^2/(2\pi)^2$ is shown in Fig. 2 labelled "isotropic scattering".

One can integrate (16) directly assuming isotropic scattering and obtain

$$\rho/\rho_0 = \sum_i \alpha_i c_i\{j_0^2(kn_i)-j_1^2(kn_i)\}. \qquad (22)$$

For face-centered-monovalent solid solutions with spherical energy surfaces $ka/2\pi=(3/2\pi)^{1/3} = 0{\cdot}78$. The corresponding value for 30 per cent α brass is $ka/2\pi = 0{\cdot}780(1{\cdot}3)^{1/3} = 0{\cdot}852$ because the number of conduction electrons is 1·5 times as large. The values of n_i are

$$n_i = \frac{a}{2}\{0, \sqrt{2}, \sqrt{4}, \sqrt{6}, \ldots\}. \qquad (23)$$

The function $\{j_0^2(kn_i)-j_1^2(kn_i)\}$ for these parameters is tabulated in Table 1.

α brass is a face-centered-cubic solid solution as is disordered $AuCu_3$. $AuCu_3$ orders below about 395° C. Long-range order has never been observed in α brass, however BIRCHENALL,[15] and later GUTTMAN,[16] used the activity coefficient for zinc

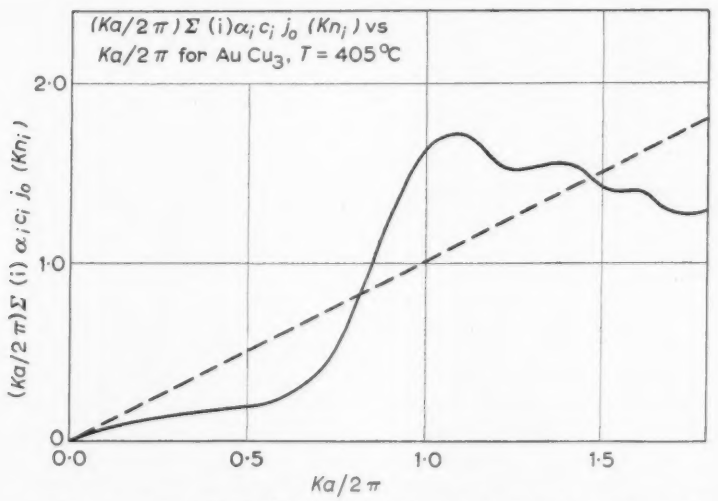


FIG. 1. The function $(Ka/2\pi)\sum_i \alpha_i c_i j_0(Kn_i)$ shown in the solid curve is calculated for a monovalent face-centered-cubic solid solution and short-range parameters α_i measured by COWLEY on $AuCu_3$ at 405°C.

in α brass to calculate that the Cu-Zn is preferred and that, therefore, ordering is possible. Probably diffusion is so slow at temperatures below the critical temperature that equilibrium order cannot be achieved in reasonable times. For an exploratory estimate of the effect of short-range order on the residual resistivity of α brass, short-range order was assumed to be the same as that measured for $AuCu_3$ by COWLEY. KEATING[16] concluded from X-ray and neutron diffraction studies that α brass at 190° C did not exhibit positive evidence of short-range order,* but it could possess as much short-range order as $AuCu_3$ at 550° C. Using values of α_i measured by COWLEY on $AuCu_3$ at this temperature one obtains from eq. (22)

$$\rho/\rho_0 = 1{\cdot}004 \qquad (24)$$

This is in poor agreement with experimental results of DAMASK[6] that at about 200° C the effect of short-range order on 30 per cent α brass is to decrease the residual resistivity about 12 per cent from that of the cold-worked material which is presumably disordered.

One should not be surprised that the calculation for $AuCu_3$ agrees with experiment better than the calculation for α brass since the assumption of spherical energy surfaces is much better in the former case. The reason is that as Zn is alloyed with Cu the number of conduction electrons is increased and the Fermi surface moves toward the Brillouin zone boundary. In 30 per cent α brass a spherical surface would approach very close to the zone boundary in the (111) direction.

VOL. 1 1956/57

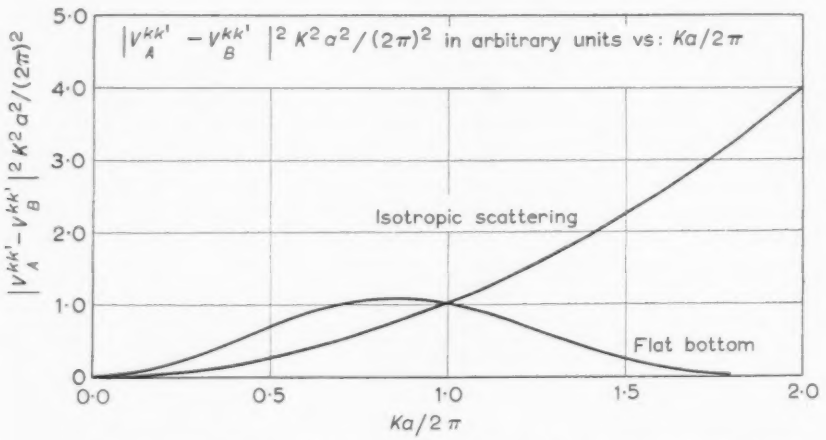


FIG. 2. The function $|V_A^{kk'} - V_B^{kk'}|^2 K^2 a^2/(2\pi)^2$ is shown for isotropic and flat-bottom scattering. The former is used in the calculation of the residual resistivity of solid solutions whose constituents have nearly the same atomic number (e.g. α brass). The latter is used in the calculation of the residual resistivity of face-centered-cubic monovalent solid solutions (e.g. $AuCu_3$).

In $AuCu_3$ a large part of integral (13) came from the region around $Ka/2\pi = 1$, where both $KF(K)$ (see Fig. 1) and $K^2|V_A^{kk'} - V_B^{kk'}|^2$ (see Fig. 2, "flat-bottomed") are large. α brass is different from $AuCu_3$ in two respects. The scattering is more nearly isotropic (see Fig. 2, isotropic scattering) and thus the function $K^2|V_A^{kk'} - V_B^{kk'}|^2$ is largest for large K. For $Ka/2\pi$ larger than 1·5, $KF(K)$ falls below K (see Fig. 1) and therefore integral (13) is reduced. Second, the increased concentration of conduction electrons due the Zn in α brass increases the radius of the Fermi sphere and consequently the upper limit of integral (13). This upper limit is the region in which $KF(K)$ is smaller than K so that integral (13) is reduced relative to its value for $F(K) = 1$ (complete disorder). Thus the change in the scattering function and the increased

* Because of the near equality of scattering powers of copper and zinc, diffraction measurements are not very sensitive to the degree of order in this system.

radius of the Fermi sphere both decrease the residual resistivity due to short-range order in α brass from that of a monovalent solid solution such as $AuCu_3$. The calculated reduction is about one-half that observed.

It should be pointed out that these calculations are crude. Both the potential and wave functions in a solid solution have been grossly approximated. The sums in (21) and (22) converge rather slowly so that cutting off the sum at the tenth shell may not be justified.

One result of this work is to point out similarity between the effect of short-range order on electronic scattering and X-ray scattering. A qualitative feel for the effect may be had by noting that residual resistance depends on the integral of the product of two functions such as shown on Figs. 1 and 2. The size and sign of the effect will depend sensitively on the details of these functions.

SUMMARY

The ratio of the residual resistivity of $AuCu_3$ quenched at 405° C, 460° C and 550° C to the residual resistivity of disordered $AuCu_3$ was computed to be 1·177, 1·167, and 1·127 respectively.

The ratio of the residual resistivity of 30 per cent α brass quenched at 190° to the residual resistivity of disordered 30 per cent α brass was computed to be 1·004.

Acknowledgement—The author wishes to thank Dr. G. H. VINEYARD for several valuable discussions.

APPENDIX A

A variational expression for the electrical resistance due to KOHLER,[9] and discussed by WILSON,[17] is as follows:

$$\rho = (\phi_i,\phi_i)/e^2\left[\frac{1}{4\pi^3}\int \phi_i v_i \frac{\partial f_0}{\partial E}\,d\boldsymbol{k}\right]^2. \qquad \text{(A1)}$$

The distribution function in $\boldsymbol{k}$ space of electrons per unit volume, $f(\boldsymbol{k})$, is related to the function Φ as follows $f = f_0 - \Phi\frac{\partial f_0}{\partial E}$ The equilibrium distribution is $f_0(\boldsymbol{k})$. The velocity $\boldsymbol{v}$ is defined as $\boldsymbol{v} = \hbar^{-1}\nabla E(\boldsymbol{k})$. The function φ_i is related to Φ in the following way:

$$\Phi = \sum_{i=1}^{3}\left[\epsilon\mathcal{E}_i + T\frac{\partial}{\partial x_i}\left(\frac{\zeta}{T}\right)\right]\phi_i,$$

where ϵ is the numerical value of the charge of the electron, $\mathcal{E}$ is the electric field, T the temperature, and ζ the Fermi energy.

The collision operator is defined as follows

$$L(\Phi) \equiv \frac{1}{8\pi^3 kT}\int V(\boldsymbol{k},\boldsymbol{k}')\{\Phi(\boldsymbol{k})-\Phi(\boldsymbol{k}')\}\,d\boldsymbol{k}', \qquad \text{(A2)}$$

and the quantity (G,G) as follows:

$$(G,G) = 1/4\pi^3\int GL(G)\,d\boldsymbol{k}. \qquad \text{(A3)}$$

The symmetric function $V(\boldsymbol{k},\boldsymbol{k}')$ is defined as follows

$$V(\boldsymbol{k},\boldsymbol{k}') = W(\boldsymbol{k},\boldsymbol{k}')f_0(\boldsymbol{k})\{1-f_0(\boldsymbol{k}')\}, \qquad \text{(A4)}$$

where $W(\boldsymbol{k},\boldsymbol{k}')$ is the probability per unit time of scattering from state $\boldsymbol{k}$ to $\boldsymbol{k}'$. From time dependent perturbation theory this function may be evaluated as follows[18]:

$$W(\boldsymbol{k},\boldsymbol{k}') = (2\pi/\hbar)|V'_{\boldsymbol{k}\boldsymbol{k}'}|^2\delta(E_i-E_f). \qquad \text{(A5)}$$

Elastic scattering has been assumed. E_i and E_f are the initial and final energies. $|V_{\boldsymbol{k}\boldsymbol{k}'}|^2$ is the matrix element of the perturbation energy.

Notice that expression (A1) is homogeneous in φ_i. Thus, multiplicative constants in φ_i cancel. Expression (A1) is also extremal, its value being smallest for the correct value of φ_i. As WILSON[17] suggests, let us evaluate (A1) for the usual approximation for Φ, namely

$$\Phi = \text{const. } v_1, \qquad \text{(A6)}$$

assuming the electric field is in the direction of the 1 axis.

Substituting eq. (A6) in eq. (A1) one obtains

$$\rho = \frac{16\pi^6}{e^2}\,\frac{(v_1,v_1)}{\left[\int v_1{}^2\frac{\partial f_0}{\partial E}d\boldsymbol{k}\right]^2}. \qquad \text{(A7)}$$

Assuming cubic symmetry and absolute zero of temperature one obtains for the denominator of eq. (A7)

$$\int v_1{}^2\frac{\partial f_0}{\partial E}d\boldsymbol{k} = -(1/3)\int_{\zeta} v^2\,dS/|\nabla_k E(\boldsymbol{k})|, \qquad \text{(A8)}$$

since $-\frac{\partial f_0}{\partial E} = \partial(E-\zeta)$ where ζ is the Fermi energy. The integration is over the Fermi surface. This integral is the average of the square of the velocity on the Fermi level times the density of states at the Fermi level.

The numerator, shown below, can also be simplified for these assumptions.

$$(v_1,v_1) = \frac{3}{(2\pi)^6 kT}\iint v_1(\boldsymbol{k})V(\boldsymbol{k},\boldsymbol{k}')\{v_1(\boldsymbol{k})-v_1(\boldsymbol{k}')\}\,d\boldsymbol{k}d\boldsymbol{k}'. \qquad \text{(A9)}$$

Since $V(\mathbf{k},\mathbf{k}')$ is symmetric in $\mathbf{k}$ and $\mathbf{k}'$, if one interchanges $\mathbf{k}$ and $\mathbf{k}'$ in eq. (A9), adds this changed equation to (A9) and divides by 2, one obtains

$$(v_1,v_1) = (2\pi)^{-6}(kT)^{-1} \iint V(\mathbf{k},\mathbf{k}')\{v_1(\mathbf{k})-v_1(\mathbf{k}')\}^2\, d\mathbf{k}d\mathbf{k}'. \quad \text{(A10)}$$

Similarly, change 1 to 2 and 3 in eq. (A10), add the three equations and divide by 3. One obtains

$$(v_1,v_1) = (2\pi)^{-6}(kT)^{-1} \iint V(\mathbf{k},\mathbf{k}')\{\mathbf{v}(\mathbf{k})-\mathbf{v}(\mathbf{k}')\}^2\, d\mathbf{k}d\mathbf{k}'/3. \quad \text{(A11)}$$

Change the integration over $\mathbf{k}$ space into integration over a constant energy surface and integration over energy, thus

$$d\mathbf{k} = dEdS/|\nabla_k E(\mathbf{k})|. \quad \text{(A12)}$$

Integrate over E. The δ-function in eq. (A5) gives an integral over the surface with energy equal to the energy of $\mathbf{k}'$, This changes eq. (A4) to

$$V(\mathbf{k},\mathbf{k}') = W(\mathbf{k},\mathbf{k}')f_0(\mathbf{k}')\{1-f_0(\mathbf{k}')\} = kTW(\mathbf{k},\mathbf{k}')\delta(E(\mathbf{k}')-\zeta), \quad \text{(A13)}$$

since $\mathbf{k}'$ enters into $f_0(\mathbf{k}')$ through the energy and

$$f_0(\mathbf{k})\{1-f_0(\mathbf{k})\}/kT = \delta(E-\zeta). \quad \text{(A14)}$$

Substituting eq. (A13) and eq. (A5) in eq. (A11), one obtains

$$(v_1,v_1) = \frac{(2\pi)^{-5}}{3\hbar} \iint |V'_{\mathbf{k}\mathbf{k}'}|^2\{\mathbf{v}(\mathbf{k})-\mathbf{v}(\mathbf{k}')\}^2\, dSdS'/|\nabla_k E(\mathbf{k})||\nabla_{k'} E(\mathbf{k}')|, \quad \text{(A15)}$$

where the integrations are over the Fermi surface in $\mathbf{k}$ and $\mathbf{k}'$. Substituting eq. (A15) and eq. (A8) in eq. (A7) one obtains

$$\rho = \frac{3\pi}{2\hbar} \frac{\iint |V'_{\mathbf{k}\mathbf{k}'}|^2\{\mathbf{v}(\mathbf{k})-\mathbf{v}(\mathbf{R}')\}^2\, dSdS'/|\nabla_k E(\mathbf{k})|\,|\nabla_k E(\mathbf{k}')|}{[\int v^2\, dS/|\nabla_k E(\mathbf{k})|]^2}. \quad \text{(A16)}$$

Notice that in this formalism the density of states at the Fermi level roughly cancels out (exactly for spherical energy surfaces). For spherical energy surface eq. (A16) becomes

$$\rho = \frac{3\pi}{\hbar(4\pi)^2} v^2 \iint |V'_{\mathbf{k}\mathbf{k}'}|^2(1-\cos\theta)\, d\Omega d\Omega', \quad \text{(A17)}$$

where $d\Omega$ is the element of solid angle, $d\Omega = ds/k^2$.

REFERENCES

1. Nordheim L. *Ann. Physik* **9,** 607, 641 (1931).
2. Mott N. F. *Proc. Cambridge Phil. Soc.* **32,** 281 (1936).
3. Muto T. *Sci. Papers Inst. Phys. Chem. Research (Tokyo)* **30,** 99 (1936); **34,** 377 (1938); Slater J. C. *Phys. Rev.* **84,** 179 (1951).
4. Murakami T. *J. Phys. Soc. Japan* **8,** 458 (1953).
5. Krivoglaz M. A. and Matysina Z. A. *Soviet Physics JETP* **1,** 103 (1955).
6. Damask A. *J. Appl. Phys.* **27,** 610 (1956); *J. Phys. Chem. Solids* **1,** 23, (1956).
7. Houston W. V. *Phys. Rev.* **51,** 184 (1940).
8. Cowley J. M. *J. Appl. Phys.* **21,** 24 (1950); *Phys. Rev.* **77,** 669 (1950).
9. Kohler M. *Z. Phys.* **124,** 772 (1948); **125,** 679 (1949).
10. Rosenthal J. E. and Murphy G. M. *Rev. Mod. Phys.* **8,** 317 (1936). For further discussion see Eyring, Walter and Kimball, *Quantum Chemistry* (John Wiley and Sons, New York, 1944) 187.
11. Schiff L. I. *Quantum Mechanics* (McGraw-Hill, New York, 1949) 77.
12. Warren B. E., Averbach B. L. and Roberts B. W. *J. Appl. Phys.* **22,** 1493 (1954).
13. Seitz F. *The Modern Theory of Solids,* (McGraw-Hill, New York, 1940) 543.
14. Birchenall C. E. *Trans. Amer. Inst. Min. Met. Eng.* **171,** 166 (1947).
15. Guttman L. *Trans. Amer. Inst. Min. Met. Eng.* **175,** 178 (1948).
16. Keating D. *Acta Met.* **2,** 885 (1954).
17. Wilson A. H. *The Theory of Metals,* (Cambridge University Press, Cambridge, 1953) 322.
18. Seitz F. op. cit., 521.

VOL. 1 1956/5

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 35–38.

THE ELECTRONIC PROPERTIES OF NICKEL-PALLADIUM ALLOYS

E. P. WOHLFARTH

Department of Mathematics, Imperial College of Science and Technology, London, England

(*Received* 23 *April* 1956)

Abstract—The magnetic properties of Ni-Pd alloys are discussed on the basis of the collective electron treatment of ferromagnetism. A series of further experimental investigations is suggested, including measurements of the specific heat at low temperatures, of the magnetization below the Curie point and in the presence of a strong magnetic field at low temperatures, of the composition and temperature variation of the electrical resistivity and of the effect of strong compression on the Curie temperature of the Pd-rich alloys. In all cases it is possible to make qualitative or quantitative predictions as to the outcome of the investigations. It is suggested that compression of pure Pd by means of practicable pressures may have the effect of rendering the metal ferromagnetic with a low Curie point.

1. INTRODUCTION

SCHINDLER, SMITH, and SALKOVITZ (private communication, cf. reference 1) have recently started an experimental investigation of the electronic properties of Ni–Pd alloys. It may be worth while, therefore, to give an account of the theoretical conclusions regarding the properties of this alloy system which may be drawn from an application of the collective electron treatment of ferromagnetism. A part of such an analysis has been very briefly summarized in earlier papers[2, 3, 4], but a complete account, including more recent work, has not so far been published. Furthermore, it is now appropriate to suggest a series of experimental investigations which may help to throw further light on the electronic structure of this alloy system in particular, and on that of ferromagnetic alloys in general.

The Ni–Pd alloys form a noninterrupted series of solid solutions having face-centred cubic structure over the whole composition range. There is strong evidence[2] that q_d, the number of holes per atom in the d band, has about the same value for Pd as for Ni, implying a similar equality for the number of conduction electrons per atom of the two metals. It is, therefore, reasonable to suppose that these numbers remain unchanged on alloying. A theoretical discussion of the electronic properties of these alloys is thus very much simpler than for nearly any other alloy system. Such a simplification arises in a two-band treatment of electrical properties[1]. In discussing magnetic and thermal properties it is reasonable to assume further that the density of states curves for the d bands near their upper ends, here assumed to have parabolic shape, remain unchanged for all compositions, implying constancy of the parameter ϵ_{0d} (the energy width of the unoccupied part of the d band in the absence of spontaneous magnetization). For a discussion on the basis of the collective electron treatment[3], therefore, the only one of the three parameters q_d, ϵ_{0d} and θ' which may be taken to vary significantly as a function of Pd concentration is the characteristic temperature θ', which gives a measure of the energy of interaction (exchange and correlation) between the d electrons.

2. ANALYSIS OF MAGNETIC DATA

The method of analysing the available magnetic data is as follows: The measurements of SADRON[5] and WENT[6] give the specific saturation magnetization σ_0 for Pd concentrations c between 0 and 70 per cent (atomic), extrapolation to higher c being facilitated by the results of other magnetic measurements[7, 8, 9] that σ_0 vanishes at about 97 per cent Pd. Now

$$\sigma_0 = q_d \zeta_0 M_B / A(c), \qquad (1)$$

where M_B is the Bohr magneton per mole and $A(c)$ the atomic weight of an alloy with Pd concentration c. Also ζ_0 is the relative magnetization at 0°K, i.e. the ratio of the resultant spin to the maximum possible if spin alignment is complete; a value $\zeta_0 < 1$ means that the d subbands with both + and − spin contain holes and so contribute, with opposite sign, to the resultant magnetization, while $\zeta_0 = 1$ means that the subband with − spin, say, is completely filled at 0°K. In (1), $A(c)$ is given

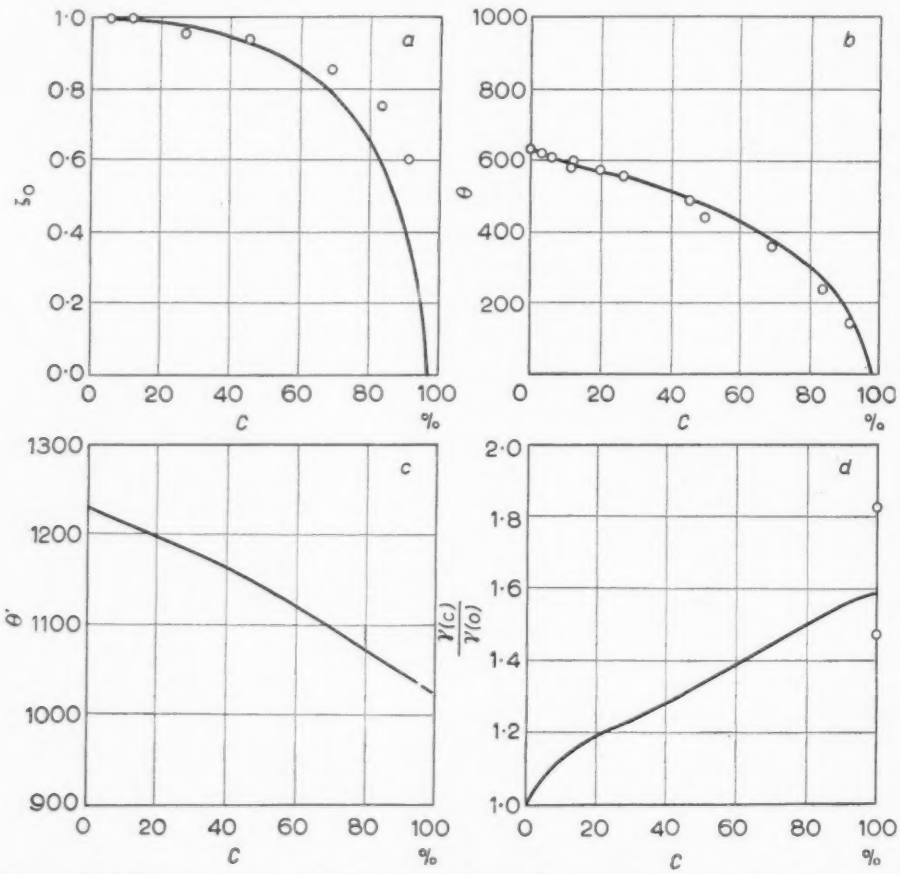


FIG. 1. (a) Full curve: values of ζ_0 from observed saturation magnetization; circles: from paramagnetic susceptibility. (b) Full curve: Curie temperatures from ζ_0; circles: observed. (c) θ' values from ζ_0. (d) Full curve: calculated ratio of electronic heat coefficients; circles: two corrected observed values for pure Pd.

by an additive law,

$$A(c) = A(0)(1+0{\cdot}818c), \qquad (2)$$

so that, if q_d is independent of c and $\zeta_0 = 1$ for pure Ni[10],

$$\sigma_0(c) = \zeta_0 \sigma_0(0)(1+0{\cdot}818c)^{-1}. \qquad (3)$$

The calculated values of ζ_0, using (3), are shown in Fig. 1(a).

Quite an independent check on these values of ζ_0 may be obtained as follows: The measurements of MANDERS[7] give the paramagnetic mass susceptibility χ of the alloys above their Curie points for c up to 98 per cent. If a family of curves is

drawn relating for each alloy $(\mu_B\sigma_0/k\theta)\chi^{-1}$ with T/θ, where θ is the Curie temperature[6, 8, 9], then approximate values of ζ_0 may be deduced by comparison with the theoretical curves[10] drawn for $0 < \zeta_0 \leqslant 1$. These are uncertain to the extent that the temperature excitation of d electrons into the conduction band[10] has been neglected. The ζ_0 values obtained in this way are shown in Fig. 1(a) by circles. An increasingly rapid lowering of ζ_0 is obtained from both sets of data, although there are some deviations between them at the higher c. In what follows ζ_0 values are taken from the full curve, i.e. as calculated from the magnetization data.

For a parabolic shape of the head of the d band STONER's calculations[11] give an implicit relation between ζ_0 and the ratio $k\theta/\epsilon_{0d}$. Hence, taking ϵ_{0d} independent of Pd concentration c and the Curie temperature of Ni as 631°K, the values of θ for any value of c may be calculated. The curve relating θ and c is shown in Fig. 1(b). The directly measured values of θ[6, 8, 9], shown by circles, are in excellent agreement with this curve, even in passing through a slight inflection at about 25 per cent Pd.

Again, the simple relation[11] between ζ_0 and the ratio $k\theta'/\epsilon_{0d}$,

$$k\theta'/\epsilon_{0d} = [(1+\zeta_0)^{2/3}-(1-\zeta_0)^{2/3}]/2\zeta_0, \quad (4)$$

makes it possible to calculate the variation of θ' with c, assuming ϵ_{0d} constant and taking[10] θ' for Ni equal to $1{\cdot}23\times10^3$ °K. The curve relating θ' with c is shown in Fig. 1(c) up to 95 per cent Pd. Extrapolation gives for pure Pd $\theta' = 1{\cdot}02\times10^3$ °K, in good agreement with the values deduced[2] from the paramagnetic susceptibility and its temperature variation at temperatures removed from the mysterious susceptibility maximum at 80°K[11a]. As shown earlier[3] by a simple argument, θ' may be taken to depend on the product of q_d and a complicated function of the interatomic distance R. The small decrease of θ' with increasing Pd concentration may in this case (q_d constant) be ascribed solely to the increase of R in passing from Ni ($R = 2{\cdot}49$ Å) to Pd ($R = 2{\cdot}75$ Å). The value of the ratio $k\theta'/\epsilon_{0d}$ for Pd deduced in this way is 0·65, only slightly less than the critical value (2/3 for parabolic band shape) for which ferromagnetism occurs.

3. DEDUCTIONS FROM THE ANALYSIS—FURTHER EXPERIMENTS

The principal result of section 2 is that alloying Ni with Pd causes a decrease of the ratio ζ_0 as shown in Fig. 1. The presence of holes in both + and − spin subbands implies that both make a contribution to the low-temperature linear electronic heat coefficient γ. With parabolic band shape the ratio of γ for an alloy with Pd concentration c to its value for Ni is given by[12]

$$\gamma(c)/\gamma(0) = 2^{-1/3}[(1+\zeta_0)^{1/3}+(1-\zeta_0)^{1/3}]. \quad (5)$$

The calculated values of this ratio are shown in Fig. 1(d). This curve represents a prediction of the low temperature thermal behavior of Ni–Pd alloys; the slight inflection at about 25 per cent Pd, analogous to that shown in Fig. 1(b), may be significant and the whole curve could be checked experimentally. At the moment only the values for pure Ni and pure Pd are available. Taking[13] $\gamma(0) = 17{\cdot}4\times10^{-4}$ cal mol^{-1}deg^{-2}, and making small corrections for the contributions from the conduction electrons and the thermal scattering of d electrons into the conduction band[10], the value of γ for Pd becomes $27{\cdot}5\times10^{-4}$. Two values, shown by circles in Fig. 1(d) (after making similar small corrections) have been obtained experimentally, PICKARD and SIMON[14] giving $31{\cdot}0\times10^{-4}$ and RAYNE[15] $25{\cdot}6\times10^{-4}$.

Another type of experiment which could in principle verify that $\zeta_0 < 1$ is to apply a strong magnetic field at low temperatures to an alloy with relatively large Pd content so that ζ_0 is indeed quite low. This should cause an increase in the effective value of ζ_0, i.e. of the low-temperature magnetization, above its value in the absence of the field by transferring electrons from the + spin to the − spin band. Consideration of this effect in the case of Ni–Cu alloys, where values of ζ_0 considerably less than 1 were also predicted[10], indicates that a significant increase of magnetization may be produced in an alloy with a relatively low ζ_0 by fields of the order 5×10^4 oersted, which are now quite practicable. Ni–Pd alloys are in many ways more suitable for this type of experiment than Ni–Cu alloys, for which an investigation of the effect gave rather inconclusive results[16].

Although the paramagnetic susceptibility above the Curie point is sensitive to the value of ζ_0, the reduced magnetization, temperature curves below

the Curie point (giving σ/σ_0 as function of T/θ, where σ is the magnetization at temperature T) are to a first approximation almost independent of ζ_0 over the significant range[11]. If, however, the parameter θ' is taken to be magnetization dependent[3, 17] then, by analogy with the magnetization curves for Ni–Cu alloys[17], it is expected that some deviations from the law of corresponding states will occur at high Pd content. Data on the magnetization of these alloys are thus desirable to see if such deviations occur. The curves are expected[8] to suffer less from tailing near the Curie point than is the case for most other Ni alloys.

The electrical resistivity and related properties of transition metals are strongly influenced by the scattering of the conduction electrons into the *d* band[18]. For Ni $\zeta_0 = 1$ and the scattering can only be into the + subband containing holes, so that it is less pronounced than for a nonferromagnetic transition metal like Pd. As shown above there is a gradual reduction of ζ_0 when these two metals are alloyed. Hence there should be a gradual *relative* increase of the resistivity ρ with increasing Pd content, since now the conduction electrons may be scattered into both subbands at all temperatures. The curve giving ρ as a function of Pd concentration at a given temperature is thus expected to be asymmetrical, the usual maximum lying nearer to the Pd-rich end. The preliminary results[1] actually show such an asymmetry, the maximum lying at 70 per cent Pd. The increase could also be detected by measuring ρ as a function of temperature and determining the difference $\Delta\rho$ between the measured value at some fairly low temperature $T < \theta$ and the value obtained by backward extrapolation of the high temperature portion of the $\rho - T$ curve to the same temperature T. The above discussion predicts that the value of $\Delta\rho$ at a given T, suitably reduced, decreases in parallel with the ζ_0 change as the Pd content increases. The lowering of ζ_0 will influence other electrical properties as well, for example magneto-resistance[19]. If $\zeta_0 < 1$ then $(1/\rho)(\partial\rho/\partial H)_T$, which is proportional to $(\partial\sigma/\partial H)_T$, remains finite as $T \to 0$ and SMIT actually observed this for some Ni–Fe and Ni–Cu alloys. No such measurements are available for Ni–Pd alloys.

As shown in Fig. 1(b) the Curie temperature θ varies rapidly with composition for the Pd-rich alloys, this variation being ascribed above to the dependence on interatomic distance R of the interaction parameter θ'. It is hence expected that if decreases of R are produced not by adding Ni to Pd but by compression, θ will turn out to be rather pressure sensitive in this composition range. It may even be possible to compress pure Pd sufficiently to make it ferromagnetic at low temperatures. Assuming R to vary linearly with composition it is found that for Pd the compression $|\Delta V/V|$ necessary to simulate an alloy with 5 per cent Ni (which is certainly ferromagnetic with a Curie temperature about 50°K) is equal to 0·0142; BRIDGMAN[20] obtained experimentally a compression 0·0152 at a pressure 3×10^4 kg cm^{-2}. He also found, however, that the compression curve has a discontinuity in slope at a pressure $1{\cdot}6 \times 10^4$. It may be that the above effects will be masked, but they may also be enhanced, by a distortion or relative shift of the energy bands. It seems, nevertheless, worth while to see if application of practicable pressures of the order mentioned will produce ferromagnetic palladium with a low Curie point.

VOL. 1 1956/5

REFERENCES

1. SCHINDLER A. I., SMITH R. J. and SALKOVITZ E. I. *Bull. Amer. Phys. Soc.* **1(3),** 114 (1956).
2. WOHLFARTH E. P. *Proc. Leeds Phil. Soc.* **5,** 89 (1949).
3. WOHLFARTH E. P. *Rev. Mod. Phys.* **25,** 211 (1953).
4. WOHLFARTH E. P. *Phil. Mag.* **45,** 647 (1954).
5. SADRON C. *Ann. Phys.* **17,** 371 (1932).
6. WENT J. J. *Physica* **17,** 98 (1951).
7. MANDERS C. *Ann. Phys.* **5,** 167 (1937).
8. MARIAN V. *Ann. Phys.* **7,** 459 (1937).
9. NÉEL L. *Ann. Phys.* **18,** 5 (1932).
10. WOHLFARTH E. P. *Proc. Roy. Soc.* **A195,** 434 (1949).
11. STONER E. C. *Proc. Roy. Soc.* **A165,** 372 (1938).
11*a*. HOARE F. E. and MATTHEWS J. C. *Proc. Roy. Soc.* **A212,** 137 (1952).
12. STONER E. C. *Proc. Roy Soc.* **A169,** 339 (1939).
13. KEESOM W. H. and CLARK C. W. *Physica* **2,** 513 (1935).
14. PICKARD G. L. and SIMON F. E. *Proc. Phys. Soc. (Lond.)* **61,** 1 (1948).
15. RAYNE J. *Phys. Rev.* **95,** 1428 (1954).
16. ARROTT A. Thesis, Carnegie Institute of Technology (1954).
17. HUNT K. L. *Proc. Roy. Soc.* **A216,** 103 (1953).
18. MOTT N. F. *Proc. Roy. Soc.* **A153,** 699 (1936); **A156,** 368 (1936).
19. SMIT J. *Physica* **17,** 612 (1951); Thesis, Leiden (1956).
20. BRIDGMAN P. W. *Proc. Amer. Acad. Arts Sci.* **77,** 187 (1949).

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 39–41.

PRELIMINARY ELECTRICAL-RESISTIVITY MEASUREMENTS OF THE NICKEL-PALLADIUM ALLOY SYSTEM

A. I. SCHINDLER, R. J. SMITH, and E. I. SALKOVITZ
Metallurgy Division, Naval Research Laboratory, Washington, D.C.

(*Received* 11 *June* 1956)

Abstract—Electrical resistivity measurements have been made on a series of nickel-palladium alloys. The resistivity reaches a maximum at approximately 70 atomic per cent palladium at 4·2°K, 77°K, and 300°K. The shift of the maximum from the 50–50 composition appears to be due to an *s–d* scattering component which increases with palladium content and persists down to liquid-helium temperatures.

INTRODUCTION

THE electronic properties of the nickel-palladium alloys have been discussed in a paper by WOHLFARTH published elsewhere in this *Journal*[1]. He found that the Curie temperature variation with composition can be explained by assuming that q_d, the number of holes per atom in the d band, remains constant at 0·6 if complete quenching of orbital angular momentum is assumed. From magnetic data he also calculated ζ_0, the relative magnetization at 0°K, and found that it varies from 1 for pure Ni to 0 for 97 per cent Pd. Since q_d remains constant with alloying, q_s, the number of electrons per atom in the s band, must also remain constant,

$$q_s = q_d = 0{\cdot}6 \equiv q,$$

for all the alloys. This results in an unusual simplification of the two-band formulation of transport properties. The two-band transport formulas require in general four parameters, the number of charge carriers of the two bands and their respective mobilities. The equality of q_s and q_d for the nickel-palladium alloys reduces the number of parameters by at least one, but for the thermoelectric power and magnetoresistance the number of charge carriers drops out of the equations completely.* We have therefore undertaken the measurement of the transport properties of the nickel-palladium alloy system. In this paper we shall describe the preliminary electrical-resistivity measurements which have been made.

DESCRIPTION OF EQUIPMENT AND PROCEDURE

Resistivity measurements were made on a group of nickel-palladium alloys which were vacuum melted by Baker and Company of Newark, New Jersey. The nickel used in the alloys was only 99 per cent pure containing traces of copper, iron, cobalt and molybdenum. The palladium used was 99·9 per cent pure. The pure nickel sample was fabricated from a spectrographic rod obtained from the International Nickel Company and was 99·99 per cent pure. Except for the pure nickel sample, the samples were in the shape of strips approximately 1 cm × 0·1 cm × 42 cm. The pure nickel sample was a rod 2·0 mm in diameter by 16·3 cm long. The samples were vacuum annealed at 800°C for 2 hours and gradually cooled for 24 hours before the measurements. The measurements were made using a modified Kelvin bridge in which errors due to lead resistance could be eliminated. With this technique, resistance could be measured to one part in 10,000. Four temperature ranges were studied, and the temperatures were measured using copper-constantan thermocouples. For helium measurements a double dewar was constructed, and the samples were suspended in the inner dewar and immersed in liquid helium. For liquid-nitrogen temperature measurements, the inner dewar was filled with liquid nitrogen. By pumping on the liquid, it was possible to vary the temperature from +64°K to +77°K. A third temperature range of −70°C to room temperature was obtained by immersing the sample in a solution of dry ice and alcohol. Finally, measurements were made between room temperature and 50°C by immersing the samples in a bath of transformer oil. The variation in temperature along the length of the samples was kept less than 0·15°C.

* WILSON[2] has shown that this simplification is found for the case of a divalent metal with overlap. Since the first zone can contain two electrons per atom, the number of occupied states in the second zone is equal to the number of vacant states in the first zone. The same behavior is true for the s and d band of the entire nickel-palladium alloy system.

EXPERIMENTAL RESULTS

The experimental results of resistivity versus temperature for all the samples thus far measured are shown in Fig. 1. For the sake of clarity, only a few of the points have been plotted. The Curie temperatures of all but the 82·3 atomic per cent palladium alloy are well above the temperature range we have investigated. For this latter sample we have gone through the Curie temperature, as can be seen by the changes of curvature of the ρ versus temperature curve. In Fig. 2 is plotted the resistivity versus composition for three temperatures 4·2°K, 77°K, and 300°K. The scatter is not due to experimental errors, but rather to spurious impurities in the various samples since for any particular sample the deviations from the continuous curve are of the same sign for the three different temperatures. Such behavior would be expected if the origin of the error were an impurity. All three curves are parabolic in shape and have maxima at approximately 70 atomic per cent palladium.

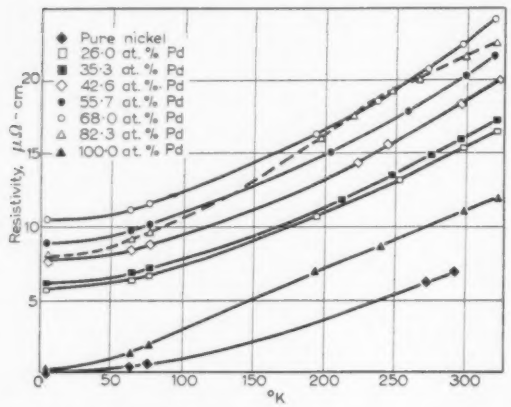


FIG. 1. The electrical resistivity as a function of temperature and composition for the nickel-palladium alloys.

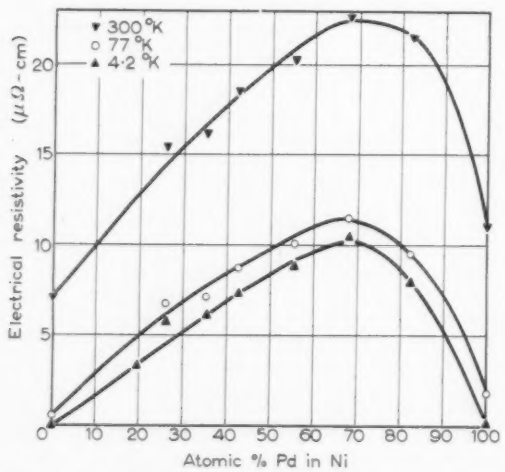


FIG. 2. The electrical resistivity as a function of composition for the nickel-palladium alloys at 4·2°, 77°, and 300°K.

DISCUSSION OF RESULTS

The electrical properties of transition metal alloys have been discussed by MOTT[3], who showed that the resistivity could be separated into two additive contributions, one arising from *s–s* scattering and one from *s–d* scattering. He concluded that because of the large effective mass of the *d*-band holes, conduction by the *d* band was negligible. His calculations showed that *s–s* scattering resulted in a Nordheim-type variation:

$$\rho \propto x(1-x),$$

where x is the fraction of the impurity atoms. The *s–d* scattering term was found to be proportional to the density of states of the *d* band at the Fermi energy. Since the number of *s* electrons remains constant with alloying, the *s–s* scattering term should result in a maximum of the resistivity at the 50–50 composition. Our data indicate that the maximum has been displaced toward the palladium-rich end for the three temperatures investigated. This displacement is probably a result of the *s–d* scattering component of the electrical resistivity. Although the number of holes in the *d* band is unaltered by alloying, the variation[1] of ζ_0 indicates that the number of holes in the two halves of the *d* band are altered with alloying. For pure nickel the holes in the *d* band are completely polarized, that is, one half of the *d* band is completely filled while all the unoccupied states are in the other half of the band. For pure palladium the holes are equally distributed in the two spin states of the *d* band. Intermediate compositions would result in a distribution of holes in the two halves of the band which can easily be calculated from the known values of ζ_0.

VOL. 1 1956/57

$$\zeta_0 = \frac{q_d{}^{\uparrow} - q_d{}^{\downarrow}}{q_d{}^{\uparrow} + q_d{}^{\downarrow}}; \quad q_d{}^{\uparrow} + q_d{}^{\downarrow} = 0{\cdot}6,$$

so,

$$q_d{}^{\uparrow} = 0{\cdot}3(1+\zeta_0),$$
$$q_d{}^{\downarrow} = 0{\cdot}3(1-\zeta_0),$$

where arrows up and down refer to the spinup and spindown halves of the *d* band. A complete description of the variation of *s–d* scattering with composition is difficult to present. However, a qualitative description can be given which is quite similar to that proposed by WOHLFARTH. For the nickel rich alloys, since one half of the *d* band is filled, only half of the *s* electrons can make *s–d* transitions without changing their spin state. As palladium is added some *d* holes of the opposite spin state become available and *s–d* scattering of both spin states of the *s* electrons can occur. Clearly, the *s–d* scattering should increase as a function of palladium content. This contribution is to be added to the *s–s* scattering term, and the sum of the two would result in a resistivity maximum which would be displaced toward the palladium-rich end. Our measurements of resistivity have a maximum at 70 atomic per cent palladium and consequently indicate that for the alloys the *s–d* scattering is an important contribution at all temperatures.

Since the alloys on which these measurements were made are rather impure for this type of work, we plan to make new measurements on purer alloy specimens. The temperature range will be extended to include the region above the Curie temperatures of all the specimens. A more complete examination will also be made in the vicinity of 4°K. At the high temperatures ζ_0 will be zero for all the alloys and consequently the *d* holes ought to be evenly distributed between the two spin states of the *d* band. The *s–d* scattering will probably not vary with composition and it seems reasonable to suppose that the maximum in resistivity will occur at the 50–50 composition. The low-temperature measurements will be made in order to determine the variation of the *s–d* scattering with temperature. We also plan to examine the other transport properties i.e., Hall effect, thermoelectric power, and magnetoresistivity over the extended temperature range for this alloy system.

Acknowledgement—We wish to thank Professor E. P. WOHLFARTH for letting us read his manuscript before publication.

REFERENCES

1. WOHLFARTH E. P. *J. Phys. Chem. Solids* **1,** 35 (1956).
2. WILSON A. H. *The Theory of Metals*, Cambridge University Press, p. 217 (1954).
3. MOTT N. F. *Proc. Phys. Soc.*, **47,** 571 (1935). *Proc. Roy. Soc.* **A153,** 699 (1936). *Proc. Roy. Soc.* **A156,** 368 (1936).

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 42–44.

THE HALL EFFECT IN THE SILVER-PALLADIUM ALLOY SYSTEM

A. I. SCHINDLER

Metallurgy Division, Naval Research Laboratory, Washington, D.C.

(*Received* 14 *June* 1956)

Abstract—Room-temperature Hall coefficient measurements have been made on the silver-palladium alloy system. The effective number of conduction electrons calculated using a one-band model is compared with that obtained for the copper-nickel alloy system and a similarity in behavior is found. The results cannot be explained using any of the multiband models proposed to date for the Hall coefficient.

VOL.
1
1956/5

INTRODUCTION

In recent years the Hall effect in ferromagnetic metals has been the subject of many investigations[1–5]. A phenomenological description relating the Hall electric field E_H, to the primary current density i, the magnetizing field H, and the intensity of magnetization M, can be written in the following manner:

$$E_H = (R_0 H + R_1 M) i \qquad (1)$$

where R_0 and R_1 have been called the ordinary and extraordinary Hall coefficient, respectively. Because of the analogous character of R_0 to the Hall coefficient R defined for nonferromagnetic metals, it has been assumed that these two coefficients are essentially the same.

An attempt to examine the variation in the number of conduction electrons with alloying can be made by measuring the Hall coefficient of a particular alloy series. Since a prerequisite for ferromagnetism is an unfilled inner band, a multiband model for the Hall coefficient is usually employed. The two-band model is the one commonly used and gives the following expression for the Hall coefficient:

$$R = -\frac{1}{Ne}\frac{(n_1\mu_1^2 - n_2\mu_2^2)}{(n_1\mu_1 + n_2\mu_2)^2} \qquad (2)$$

where N is the number of atoms per cubic centimeter, e is the electronic charge, and n and μ are the number of charge carriers per atom and their mobility; the subscripts 1 and 2 refer to electrons and holes.

In the ferromagnetic alloys the variation in the number of Bohr magnetons with alloying is usually available, and this quantity can be related to the number of holes (n_2) in the unfilled inner band. Since the total number of electrons in the two outer bands is known, a simple subtraction gives the value of n_1. The analysis now becomes one in which the values of n_1 and n_2 are inserted into eq. (2) and values of μ_1 and μ_2 are adjusted until a reasonable fit is obtained. In most cases however, the disconcerting fact is that, in order to obtain a fit for R_0, the value of μ_2 is found to be negative. To ignore this dilemma it is customary to define an effective number of conduction electrons n^* and use a one-band description of the Hall coefficient:

$$R = -\frac{1}{n^* Ne} \qquad (3)$$

The simplest alloy system for which this type of analysis has been carried out is the copper-nickel alloy system[1]. The magnetic-moment data for pure nickel indicate that nickel has 0·6 hole per atom in the 3*d* band and consequently 0·6 electron per atom in the 4*s* band (disregarding orbital angular momentum). The moment data for the nickel-copper alloys indicate that the number of holes in the 3*d* band decreases linearly as a function of copper content and goes to zero with the addition of 60 atomic per cent copper. This implies that the number of 4*s* electrons per atom remains relatively unaltered as copper is alloyed with nickel until the copper content reaches 60 atomic

per cent. As additional copper is added, the number of 4*s* electrons per atom should increase linearly from 0·6 to 1. However, analysis of R_0 for the nickel-copper alloy system indicates that n^* has a minimum value of 0·3 to 0·4 at 70 atomic per cent nickel, 30 atomic per cent copper, which is off by a factor of 2 from the values obtained with magnetic-moment data.

To determine whether the separation of the two Hall coefficients is complete, it would be desirable to repeat the Hall measurements on a similar but nonferromagnetic alloy system. Such a system is the silver-palladium alloy system.* It has been shown by MOTT[6] that the variation with composition of the number of 4*d* and 5*s* electrons of this system is the same as the variation of the number of 3*d* and 4*s* electrons for the copper-nickel system. Both alloy systems have the face-centered cubic structure and are completely soluble at all temperatures.

DESCRIPTION AND RESULTS OF MEASUREMENTS

Room-temperature direct-current measurements were therefore carried out on a series of silver-palladium alloys. The sample dimensions were $5\times1\times0{\cdot}1$ centimeters and the error in the thickness was $\pm 1/2$ per cent. Since the samples are nonferromagnetic, all measurements were made with magnetic fields varying from 0 to 12 kilogauss and current densities up to 120 amperes per square centimeter. No attempt was made to correct for the Ettingshausen effect, since a rough calculation showed that such a correction was negligible.

The results obtained for the Hall coefficient as a function of composition appear in Fig. 1. The sign of the Hall coefficient is negative for all the alloys, thus indicating conduction primarily by electrons. The absolute value of the Hall coefficient rises almost linearly with composition as silver is added to palladium, reaches a maximum at 44·2 atomic per cent silver, and then falls off as further silver is added. Values obtained by WORTMANN are also shown, and the divergence is seen to be greatest near the 50–50 composition.

* Hall measurements on the silver-palladium system had already been carried out by J. WORTMANN *Ann. Physik* **18,** 233 (1933), but, because of the size of his samples, his accuracy was quite low. The results reported here differ considerably from his.

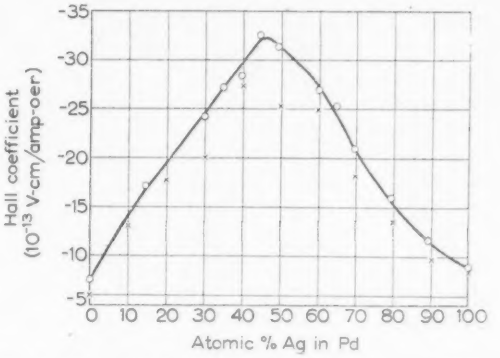


FIG. 1. Hall coefficient as a function of alloy composition. × values obtained by J. WORTMANN.

Calculations have been made of n^*, and the results appear in Fig. 2. It is noted that n^* decreases as a function of silver content and reaches a minimum value of 0·3 electron per atom at 50 atomic per cent silver and is again off by a factor of 2.

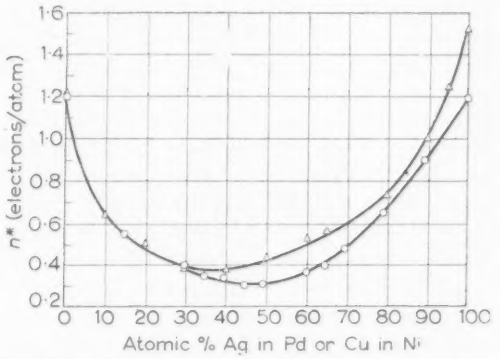


FIG. 2. Effective number of electrons per atom as a function of alloy composition.
△ = Cu–Ni; ○ = Ag–Pd.

The values of n^* obtained from the room-temperature values of the Hall coefficient for the copper-nickel system are also plotted. The similar behavior of n^* for these two alloy systems is apparent.

DISCUSSION OF RESULTS

The silver-palladium system was chosen for this investigation not only because of its similarity to the copper-nickel system, but also because it has been shown[7] that transport properties of the silver-palladium alloys are more easily understood. Electrical resistivity and magnetic susceptibility

measurements indicate that the holes in the 4*d* band are completely filled at 60 atomic per cent silver, but holes in the *d* band of the copper-nickel alloys seem to persist well beyond 60 atomic per cent copper. One would expect that the effective number of conduction electrons calculated from the Hall coefficient would remain constant at 0·6 up to 60 atomic per cent silver and then increase linearly to 1 for pure silver. As we have seen, however, the variation of n^* versus composition does not even approximate this behavior. For the alloys containing more than 40 atomic per cent palladium, a two-band Hall coefficient model ought to be applicable. Using eq. (2) and the Hall coefficient data for pure palladium, the ratio of mobilities of the *d* to *s* band is found to be 0·33 which is similar to the value obtained for pure nickel. However, the two-band model again breaks down rather quickly when the values of n^* become less than 0·6.

In a recent paper, PUGH[8] was able to explain the low values of n^* for the copper-nickel alloys by using a four-band Hall coefficient model. He divided both the 3*d* and 4*s* band into two subbands, one with spin parallel and the other with spin antiparallel. Since most of the nickel-copper alloys are ferromagnetic, the holes in the *d* band are polarized, that is, one half of the 3*d* band is completely filled while all of the holes in *d* band are in the antiparallel half of the band. He assumed that for the samples having a low value of n^*, only *s*-band conduction is important. And further, since one half of the *d* band is filled, half of the *s* electrons can make *s*–*d*, as well as *s*–*s* transitions. This decreases their mobility so that half of the *s* electrons make a contribution to conduction which is negligible. In this way, PUGH's calculations indicated that values of n^* of 0·3 were not unreasonable. The silver-palladium alloys are nonferromagnetic however, therefore the holes in the *d* band are not polarized. The four-band model, then, would give essentially the same result as the two-band model and cannot be used to explain the low values of n^* found for the silver-palladium alloys.

Since there is remarkable similarity in the values of n^* for both the nickel-copper and silver-palladium system at the nickel- and palladium-rich end, it appears that R_0 of eq. (1) is the same as the ordinary Hall coefficient of nonferromagnetic metals. However, the variation of n^* with composition obtained from R_0 cannot be explained by any of the band models of the Hall effect thus far proposed. It seems obvious to us that this should be the case. All of the band models in vogue are based upon two assumptions which are not valid for transition metal alloys. The first is that the two or more bands are completely isolated from one another and no electron transitions take place between them, and their conductivities are additive. The shape of the Hall coefficient versus composition curve found by us is very similar to that found for the electrical resistivity versus composition for the same system. MOTT[6] was able to explain the resistivity variation of the silver-palladium system by assuming two types of scattering; one involving *s*–*s* transitions, the other involving *s*–*d* transitions. It seems likely that the Hall coefficient also contains a contribution resulting from interband transitions.* However, attempts to break up the Hall coefficient into two parts similar to the separation accomplished by MOTT for the electrical resistivity proved quite difficult to carry out. The second assumption is that the Fermi surfaces are either spherical or elliptical in shape. This also does not seem to be correct for transition metals. Until more complete band calculations of these alloys are made it seems hopeless to expect anything more than an order of magnitude fit.

Acknowledgement—I would like to thank Dr. E. I. SALKOVITZ for many stimulating discussions during the course of the work.

REFERENCES

1. SCHINDLER A. I. and PUGH E. M. *Phys. Rev.* **89,** 295 (1953).
2. FONER S. and PUGH E. M. *Phys. Rev.* **91,** 20 (1953).
3. JAN J. P. *Helv. Phys. Acta* **25,** 677 (1952).
4. JAN J. P. and GIJSMAN J. M. *Physica* **18,** 339 (1952).
5. SMIT J. and VOLGER J. *Phys. Rev.* **92,** 1576 (1953).
6. MOTT N. F. *Proc. Phys. Soc.* **47,** 571 (1935).
7. COLES B. R. *Proc. Phys. Soc.* **B65,** 221 (1952).
8. PUGH E. M. *Phys. Rev.* **97,** 647 (1955).

* In a recent paper, *Phys. Rev.* **101,** 1254 (1956), COLES also indicates that knowledge of the scattering mechanisms is important in the discussion of the Hall coefficients of metal alloys.

VOL. 1 1956/5

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 45–60.

ANALYSIS OF FERROMAGNETIC AND ANTIFERROMAGNETIC SECOND-ORDER TRANSITIONS

J. A. HOFMANN, A. PASKIN, K. J. TAUER, AND R. J. WEISS

Materials Research Laboratory, Ordnance Materials Research Office, Watertown, Massachusetts

(*Received* 30 *April* 1956)

Abstract—Procedures are described for separating the magnetic contribution to the specific heats of ferromagnetic and antiferromagnetic materials. In analysing antiferromagnetic salts, it is found that the specific heats of binary salts can be fitted by combinations of two Debye functions. In both ferromagnetic and antiferromagnetic substances the total energy of ordering W is determined. The ratio W/NkT_c is found to compare favorably with calculations of this quantity by the Kramers-Opechowski and Bethe-Peierls-Weiss-Li methods based on the Heisenberg exchange formulation. The exchange integral J is calculated from W. For Ni and Fe, J is smaller than the value determined from spin-wave considerations of low-temperature saturation magnetization measurements, but for Gd the two values are the same. For the antiferromagnetic substances the magnitudes of J for the isomorphic compounds are found to lie on smooth curves when plotted versus spin.

INTRODUCTION

THERE have appeared in the literature a number of of specific-heat curves for ferromagnetic and antiferromagnetic materials. These curves exhibit the lambda shapes characteristic of second-order transitions, peaking sharply at the Curie temperature T_c. There are three possible contributions to the observed specific-heat curve of any magnetic material: lattice, electronic, and magnetic. The magnetic contribution can be extracted from the observed specific-heat curve so as to obtain a specific-heat curve due purely to the magnetic interactions. This magnetic specific-heat curve leads to quantities which can be compared with theory and with the results of other magnetic measurements.

One quantity of interest is the total area under a magnetic specific-heat curve. This area is equal to the energy W required to take the spin system from the state of approximately complete magnetic order at $T \ll T_c$ to the state of approximately complete magnetic disorder at $T \gg T_c$. Another quantity of interest is the total entropy gain in going from the ordered to the disordered state. This is easily calculated from the magnetic specific-heat curve. Simple considerations show that the theoretical entropy increase in going from a completely ordered spin state to a completely disordered spin state is equal to $Nk \ln(2s+1)$ where s is the spin per atom and N is the total number of magnetic atoms. Hence, a necessary although not sufficient condition for the validity of a separated magnetic specific-heat curve is that the associated total magnetic entropy gain approximates $Nk \ln(2s+1)$.

In the present paper, magnetic specific-heat curves have been deduced for a number of ferromagnetic and antiferromagnetic materials from their total specific-heat curves. An important part of this separation in the case of the antiferromagnetic materials is the procedure developed for estimating the lattice contribution to the total specific heats of binary compounds. The resulting experimental values of W/NkT_c for the ferromagnetic and antiferromagnetic materials are compared with theoretical calculations of this quantity based on the Heisenberg exchange model. Experimental values of the exchange parameter J that appear in the Heisenberg exchange expression are determined from the experimental values of W. For the ferromagnetic materials the values of J are compared with those determined by the application of spin-wave theory to the results of measurements of low-temperature saturation magnetization. For the antiferromagnetic materials, the values of J are plotted as a function of spin, and

correlations are noted among the values of J for isomorphic compounds.

THEORY

In our analysis of magnetic specific-heat curves it is necessary to evaluate W, the change in energy in passing from a magnetically ordered state to a magnetically disordered state. At sufficiently high temperatures (relative to the Curie temperature), not only will the long-range order vanish but the short-range order will be negligible and the energy of magnetic order likewise small. Thus, W is simply equal to the magnitude of the energy of magnetic order at $T = 0°$K.

If it is assumed[1] that the interaction energy V_{ij} between the neighboring ith and jth magnetic atoms is of the form

$$V_{ij} = -2J\boldsymbol{S}_i \cdot \boldsymbol{S}_j, \tag{1}$$

where J is an exchange energy parameter and $\boldsymbol{S}$ the spin on an atom, then W can readily be evaluated theoretically. VAN VLECK[2] has shown that one can apply ferromagnetic formulations to the antiferromagnetic case provided nearest magnetic neighbors are not nearest magnetic neighbors to one another. Fortunately, this is in fact the case in the antiferromagnetic salts here studied. Therefore, in most of our subsequent work, formulae for the two types of magnetism will be represented by a single relationship with J represented by its absolute value. In both the ferromagnetic[3] and antiferromagnetic[4] cases it is found that

$$W = NZs^2|J|, \tag{2}$$

where Z is the number of nearest magnetic neighbors, N is the number of magnetic atoms in the crystal, and s is the atomic spin.

In order to compare eq. (2) with values of W obtained from specific-heat data we must somehow evaluate $|J|$. Fortunately theory enables one to express $|J|$ in terms of the Curie or Néel temperature T_c. Eq. (2) is therefore rewritten as follows

$$W/NkT_c = Zs^2|J|/kT_c, \tag{3}$$

and values of $|J|/kT_c$ are obtained from existing calculations.

As a first approximation the ordering process is treated as a long-range-order problem analagous to the Weiss molecular-field approach. Using this method, HEISENBERG[1] (ferro) and VAN VLECK[2] (antiferro) found the following result:

$$|J|/kT_c = 3/2Zs(s+1). \tag{4}$$

Better estimates of $|J|/kT_c$ can be obtained if some effects of short-range order are included. Such calculations have recently been published by BROWN and LUTTINGER[5]. The ferromagnetic ratios of $|J|/kT_c$ are taken from their calculations based on both the Kramers-Opechowski[6, 7] method and the Bethe-Peierls-Weiss method[8]. The BROWN and LUTTINGER[5] antiferromagnetic calculations are based on the Bethe-Peierls-Weiss-Li method[9]. Although calculations based on short- plus long-range order are not exact over the whole temperature range, they should be quite good in the region of the critical temperature. Thus, it is just the desired ratio of $|J|/kT_c$ which is most accurately given by these short-range-order calculations. In this paper these short- plus long-range order calculations will be referred to as cluster calculations, and long-range-order calculations as simple Heisenberg theory.

On the basis of simple Heisenberg theory the following relationship for W/NkT_c is obtained by combining eqs. (3) and (4),

$$W/NkT_c = 3s/2(s+1). \tag{5}$$

In the case of cluster calculations, W/NkT_c cannot be expressed in closed form. The results of combining eq. (3) with BROWN and LUTTINGER's values for $|J|/kT_c$ are shown graphically in Figs. 1 and 8 along with simple Heisenberg theory.

SEPARATION OF MAGNETIC SPECIFIC-HEAT CURVES

The procedure for the separation of the magnetic contribution to the total observed specific-heat curves is somewhat different for the ferromagnetic and antiferromagnetic materials considered here, and the two cases will be discussed separately.

1. *Ferromagnetic separations*

In the ferromagnetic case there are three contributions to the observed specific-heat curves: lattice, electronic, and magnetic. It is assumed that these three contributions are additive. The magnetic contribution is determined by estimating the lattice and electronic contributions and subtracting them from the observed specific-heat curve.

VOL. 1 1956/5

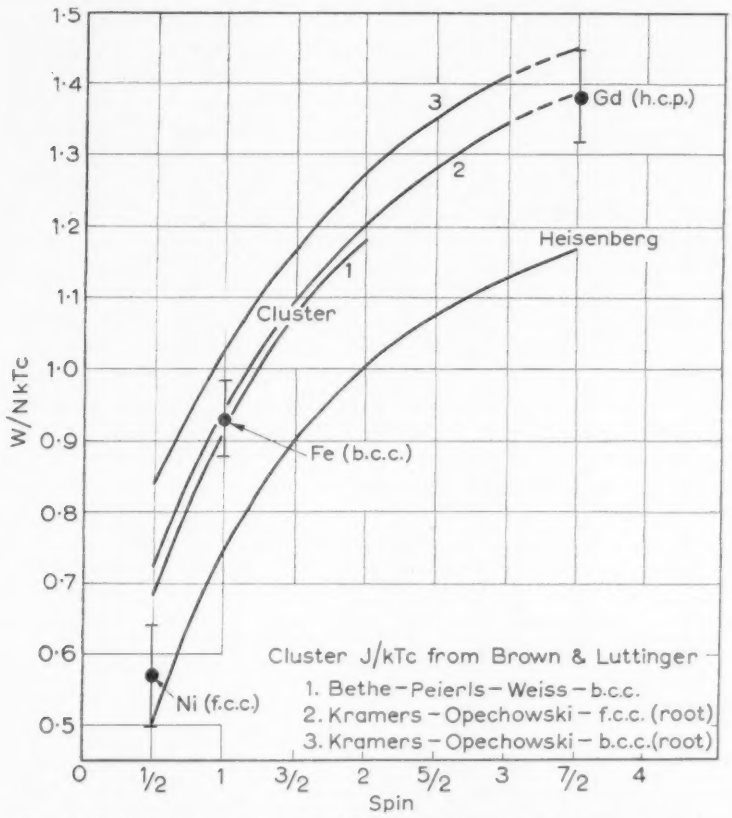


FIG. 1. Calculated values of W/NkT_c as a function of spin based on various methods using the Heisenberg exchange formulation. Included are the values of Fe, Ni, and Gd determined from their specific-heat curves.

(a) *Electronic contribution.* On the free-electron model, the electronic specific heat may be written as[10]

$$C_v(\text{el}) = \gamma T\left(1-\frac{3\pi^2}{10}\left(\frac{T}{T_0}\right)^2\right), \qquad (6)$$

where $\gamma = \pi^2 nk/2T_0$, n is the number of effective electrons, and T_0 is the degeneracy temperature. γ is determined experimentally as discussed below. T_0 then follows from the value of γ.

(b) *Lattice contribution.* In the present work on ferromagnetic substances we are concerned with pure metallic elements, and the lattice specific heats at constant volume are given to sufficient accuracy for our purposes by single Debye functions $f_D(\Theta/T)$ with constant Debye temperatures Θ.

(c) *Determination of γ and Θ.* For $T \ll \Theta$, the Debye specific heat is given by bT^3 where the constant b is a function of Θ. Furthermore at such temperatures the second term in eq. (6) is negligible since $T \ll \Theta \ll T_0$. The total specific heat at constant volume is then given by

$$C_v(\text{total}) = \gamma T + bT^3 + \text{magnetic contribution}. \qquad (7)$$

The magnetic contribution is not known until the magnetic specific-heat-curve separation has been achieved. If $T \ll T_c$, however, the magnetic contribution is small and is neglected. The usual procedure is then to determine γ and b by matching $\gamma T + bT^3$ to experimental low-temperature specific-heat curves. This is done by dividing the experimental specific heat by T, plotting the result *vs.* T^2, and extrapolating the curve to $T = 0°\text{K}$. The value

of γ is equal to the $T = 0°K$ intercept, and the value of b is equal to the slope of the curve. The value of Θ follows from the value of b. The most accurate determination of γ results from carrying out this procedure as close as possible to $T = 0°K$. This is not true, however, in the determination of Θ. What we desire is that value of Θ which leads to the most accurate lattice specific-heat curve over a temperature range of many hundreds of degrees. The procedure used in the present work consists of determining γ by matching $\gamma T+bT^3$ to experiment in the lowest temperature range possible. We then take this value of γ and determine the value of Θ which best matches the expression $\gamma T+Bf_D(\Theta/T)$ to the experimental specific-heat curve in the temperature region $T \approx \Theta/10$. In this region, the temperature is still sufficiently less than T_c to neglect the magnetic contribution. The quantity B is a correction for lattice expansion discussed below, and we use the Debye function $f_D(\Theta/T)$ rather than bT^3 since the latter expression is not sufficiently accurate for temperatures above $\Theta/15$. The values of γ and Θ so determined are used with eq. (6) and the Debye function $f_D(\Theta/T)$, respectively, to give us the electronic specific heat at constant volume and the lattice specific heat at constant volume.

(d) *Correction for constant pressure.* Since specific-heat measurements are made at constant pressure rather than constant volume, it is necessary to apply to the Debye function a correction for lattice expansion of the form

$$C_p(\text{lat}) = (1+\alpha GT)f_D(\Theta/T), \qquad (8)$$
$$= Bf_D(\Theta/T),$$

where α is the coefficient of volume expansion and G is the Grüneisen constant. (In the low-temperature range where $\gamma T+bT^3$ is matched to experimental data, the correction is negligible and bT^3 may be used directly.)

The magnetic specific-heat curve is obtained by subtracting the appropriate electronic and lattice curves from the observed total specific-heat curve.

2. *Antiferromagnetic separations*

The antiferromagnetic materials considered here are all nonmetallic, and hence they have no electronic contribution to the specific heat. They are more complicated, however, in that the lattice contribution can no longer be represented by a single Debye function. As the pertinent antiferromagnetic compounds are binary in nature, subsequent discussion will pertain to the technique of evaluating the lattice specific heats of binary compounds.

In the past it has been customary to treat the optical branch of the vibrational spectrum of such crystals by adding a sufficient number of Einstein specific-heat terms to a Debye term so that the total number of specific-heat terms equals $(n+m)$, the number of atoms in the chemical formula of the binary compound R_nX_m (this insures that the specific heat in the limit of high temperature approaches the value $(n+m)3R$). Using this

Table 1. A compilation of the two-Debye Θ's found to yield agreement between calculated and observed specific-heat curves. Ratios of Debye Θ's are also compared with the square root of the appropriate ratios of the masses of the respective atoms of these binary compounds.

Compound R_nX_m	Θ_R, Θ_X	$\frac{\Theta_X}{\Theta_R}$	$\left(\frac{M_R}{M_X}\right)^{\frac{1}{2}}$
ThO_2 [1]	250, 775	3·10	3·82
Al_2O_3 [2]	750, 1100	1·47	1·3
Ga_2O_3 [3]	360, 1075	2·98	2·1
MgF_2 [4]	375, 750	2·00	1·17
Cr_3C_2 [5]	500, 1000	2·00	2·1
HfO_2 [6]	275, 850	3·10	3·3
ZnF_2 [7]	250, 620	2·48	1·87
VN [8]	375, 1100	2·66	1·91
VC [9]	475, 1400	2·95	2·06
ZrN [10]	400, 900	2·25	2·57
CaF_2 [11]	375, 550	1·47	1·58

(1) OSBORNE D. W. and WESTRUM E. F. *J. Chem. Phys.* **21,** 1884 (1953).
(2) KERR E. C., JOHNSTON H. L., and HALLETT N. C. *J. Amer. Chem. Soc.* **72,** 4740 (1950).
(3) ADAMS G. B. and JOHNSTON H. L. *J. Amer. Chem. Soc.* **74,** 4788 (1952).
(4) TODD S. S. *J. Amer. Chem. Soc.* **71,** 4115 (1949).
(5) DESORBO W. *J. Amer. Chem. Soc.* **75,** 1825 (1953).
(6) TODD S. S. *J. Amer. Chem. Soc.* **75,** 3035 (1953).
(7) STOUT J. W. and CATALANO E. *J. Chem. Phys.* **23,** 2013 (1955).
(8) SHOMATE C. H. and KELLEY K. K. *J. Amer. Chem. Soc.* **71,** 315 (1949).
(9) SHOMATE C. H. and KELLEY K. K. *J. Amer. Chem. Soc.* **71,** 315 (1949).
(10) TODD S. S. *J. Amer. Chem. Soc.* **72,** 2914 (1950).
(11) TODD S. S. *J. Amer. Chem. Soc.* **71,** 4115 (1949).

VOL. 1 1956/5

method we have in general been unable to fit satisfactorily the specific-heat data of binary compounds over the usual temperature range. However, we have obtained remarkably good fits to specific-heat data using only Debye terms with two Debye parameters; one Debye temperature Θ for each type of atom in the binary compound. In the light of the apparent generality of our results we shall discuss in detail the two-Debye method of matching specific-heat data of binary compounds.

(a) *Generality of method.* To establish the generality of our method, eleven nonmagnetic salts were analysed in addition to ten antiferromagnetic salts. Restricting ourselves for the moment to the nonmagnetic salts, a good match to the specific-heat data was found over wide temperature ranges using the two-Debye method. The necessary Debye Θ's are listed in Table I for the compounds studied. Two experimental curves are given in Fig. 2 along with theoretical curves obtained using

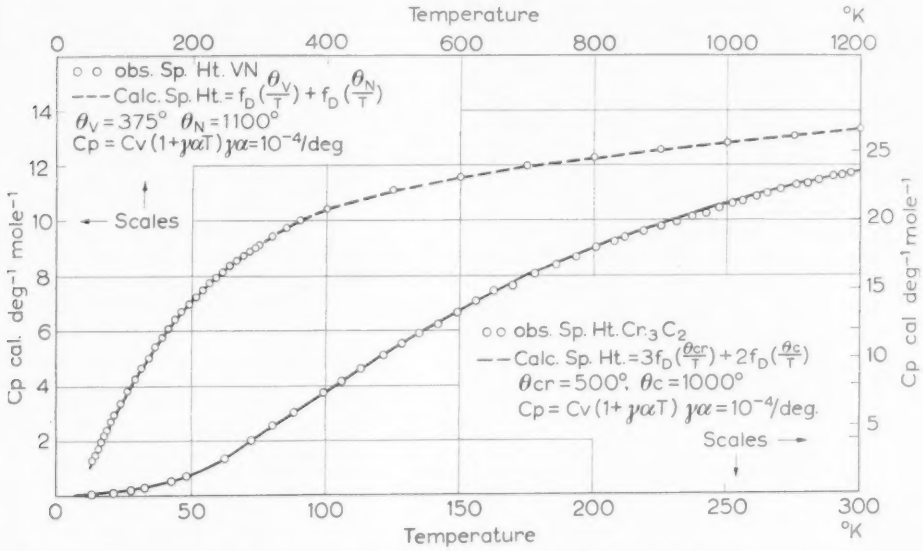


FIG. 2. Observed and calculated specific heats of VN and Cr_3C_2. Calculations are based on a two-Debye model with the Debye characteristic temperatures as indicated.

the two-Debye method. We were unable to obtain a good fit over an appreciable temperature range in these cases by using a single Debye function together with a single Einstein function. Thus, it is concluded that the two Debye method is more desirable not only in that it seems to reproduce specific-heat data of binary compounds, but that it does so requiring only two Debye parameters. Thus, the specific heat of a binary compound R_nX_m can be represented by the relationship

$$C_v = nf_D(\Theta_R/T) + mf_D(\Theta_X/T), \tag{9}$$

where $f_D(\Theta/T)$ is the familiar Debye function and Θ_R and Θ_X refer to the Debye temperatures associated with R and X, respectively.

On the basis of the following simple arguments we have been able to find a useful starting point for the determination of the two Debye Θ's required for specific-heat curve matching. In calculations pertinent to X-ray diffraction work, DEBYE has shown that $\overline{u^2}$, the mean square displacement of the atoms in a crystal, is given by

$$\overline{u^2} \propto T/M\Theta^2, \tag{10}$$

where M is the mass of the atom. Making the simplifying assumption that all the atoms in the crystal have the same mean-square displacement, it follows that

$$\Theta_R/\Theta_X = (M_X/M_R)^{\frac{1}{2}}. \tag{11}$$

The usefulness of eq. (11) as a starting point in using the two-Debye method is indicated in Table I by comparing the appropriate ratios of masses to the ratios of Debye Θ's chosen as the best fits.

(b) *Corrections for lattice expansion.* Measurements of

specific heats are usually made at constant pressure. The method of converting expressions for specific heats at constant volume to specific heats at constant pressure is similar to that outlined in the section on ferromagnetic separation. The only difference is that direct measurements of α and G for the binary compounds are generally not available. Therefore, our general approach was to assume that the product αG is equal to 10^{-4} deg^{-1}. With this assumption all but one of the nonmagnetic compounds were matched. In the case of ZnF_2, a value of $0{\cdot}5\times10^{-4}$ deg^{-1} was necessary to get a good fit in the "high-temperature" region of the curve. Fortunately, our method of entropy matching (discussed in the following section) minimizes the effect of any possible mismatch due to a poor selection of αG, for such an entropy-matching method guarantees a good match of specific heat in the lower-temperature region (see Fig. 3) where most of our antiferromagnetic subtraction was performed.

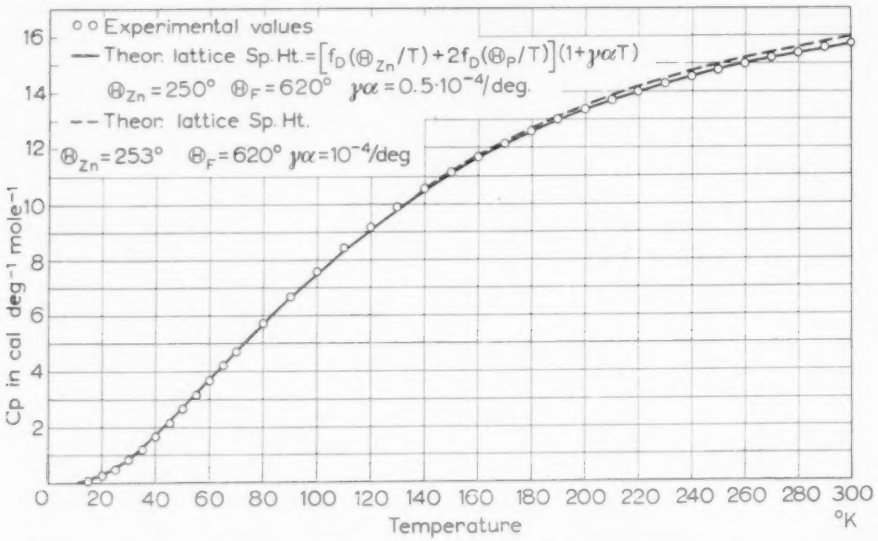

o o Experimental values
— Theor. lattice Sp. Ht. = [f_D(Θ_Zn/T) + 2f_D(Θ_F/T)](1+γαT)
Θ_Zn = 250° Θ_F = 620° γα = 0.5·10⁻⁴/deg.
-- Theor. lattice Sp. Ht.
Θ_Zn = 253° Θ_F = 620° γα = 10⁻⁴/deg
Cp in cal deg⁻¹ mole⁻¹
Temperature
°K

FIG. 3. Observed and calculated specific heats of ZnF_2. Calculations are based on a two-Debye model. The effect of varying the lattice parameters is shown.

However, in general, we do not believe that the product αG deviates from 10^{-4} deg^{-1} by more than about 50 per cent.

Should the magnetic anomaly occur at high temperature ($T_c \geq \Theta$), where the magnitude of the correction is no longer small, it is necessary to know αG more accurately. In cases where specific-heat data exist at high enough temperatures, a method is at hand to determine αG directly. Making use of the fact that for $T > \Theta$, $Cp \rightarrow 3R(n+m)(1+\alpha GT)$ for a binary compound R_nX_m, it follows that

$$\frac{1}{3R(n+m)}\frac{dC_p}{dT} \simeq \alpha G. \tag{12}$$

αG was evaluated in several cases using eq. (12), and was found to be $\sim 1{\cdot}5\times10^{-4}$. Thus, in summary, it can be seen that if the specific heat anomaly occurs at low temperature, deviations of $\alpha G = 10^{-4}$ deg^{-1} will not affect our subtraction, while for high-temperature anomalies αG can be measured directly.

(c) *Application of separation procedure.* Having found a reasonable method of matching specific heat, the question is how best to match a specific-heat curve with a magnetic anomaly. There are two methods that we have used: specific-heat matching and entropy matching.

The technique of specific-heat matching consists of matching the specific-heat data in regions removed from the anomaly, subtracting the lattice contribution, and then comparing the entropy of the magnetic curve with the theoretical value $R\ln(2s+1)$. In the event that the entropy differs from the theoretical value, other fits are made until the observed entropy is brought into agreement with the theoretical entropy value. The possibility of making other fits is due to the fact that in the specific-heat-matching technique there is some ambiguity as to where the magnetic contribution vanishes and a pure lattice specific heat exists. The entropy consistency requirement helps in objectively determining the points at which to

match the two-Debye lattice specific heat to the total specific-heat curve.

In entropy matching, the entropy requirement is built into the matching technique. At temperatures much greater than the Curie temperature, the spins have become approximately disordered and the magnetic entropy is $R\ln(2s+1)$. The lattice-entropy curve in this temperature range is therefore obtained by subtracting $R\ln(2s+1)$ from the experimental entropy curve. Two Debye temperatures Θ are determined by matching to the lattice-entropy curve a theoretical entropy curve made up of two Debye entropy functions. These Θ's are then used to construct a lattice specific-heat curve from eq. (9). An attractive feature of such a procedure is that the magnetic entropy is guaranteed to be $R\ln(2s+1)$, and the consistency requirement is thus an inherent part of the matching technique. This method is therefore the more desirable procedure of matching specific-heat curves provided the data extend to high-enough temperatures. As an illustration of this method consider the NiF_2 specific-heat curve of STOUT and CATALANO[11] shown in Fig. 4. Table 2 illustrates

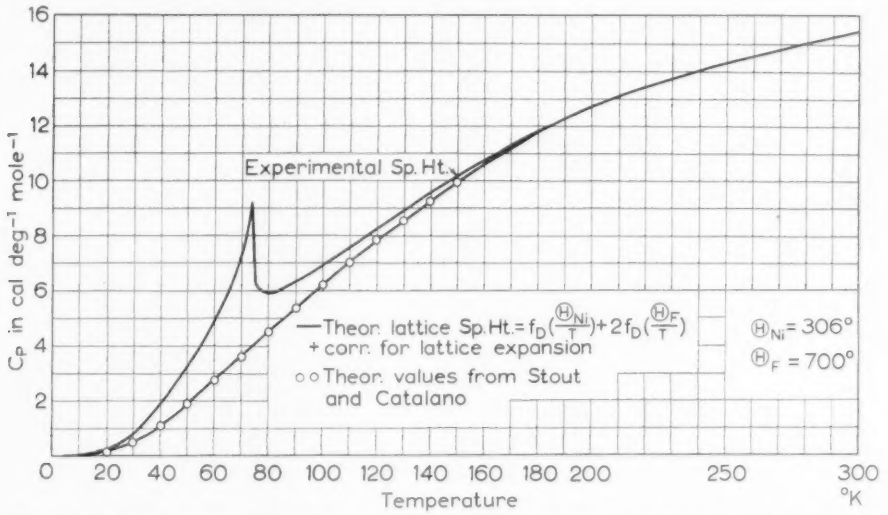


FIG. 4. Observed total specific heat of NiF_2 compared with calculated lattice specific heats based on the two-Debye model and STOUT and CATALANO's corresponding-state calculations.

Table 2. An illustration of the sensitivity of the entropy-matching method to the choice of Debye characteristic temperatures Θ_{Ni} and Θ_F for NiF_2. The experimental entropies $Sv_{(exp)}$ are given by $\int(C_v/T)dT - R\ln(2S+1)$ and the calculated entropies S_{calc} are obtained from two-Debye-method calculations.

		S_{calc}		
T°K	$Sv_{(exp)}$	$\Theta_{Ni}=300$ $\Theta_F=705$	$\Theta_{Ni}=306$ $\Theta_F=700$	$\Theta_{Ni}=300$ $\Theta_F=700$
300	15·23	15·30	15·28	15·39
250	12·61	12·64	12·59	12·70
200	9·69	9·75	9·73	9·79

the sensitivity of the entropy-matching technique to small changes in the Debye characteristic temperatures. It can also be seen that the two-Debye method matches the nonmagnetic entropy to within one per cent.

It is perhaps of further interest to note that our theoretical lattice curve in the region of magnetic anomaly matches the theoretical lattice curve of STOUT and CATALANO[11] obtained by "corresponding state" considerations. Although at first the two methods seem quite different, it can be shown [12] that they are identical if the ratio of the two Debye Θ's is a constant for the materials being compared. This seems to be the case for the isomorphic fluorides studied by STOUT and CATALANO.[11] It can also be shown[12] that corresponding-state considerations when applicable should be applied to specific-heat data converted from constant pressure to constant volume. In the case of the fluorides, the

results of the two-Debye method and corresponding-state method using data at constant volume are in excellent agreement. The corresponding-state method seems, however, to be invalid when isomorphic salts of considerably different masses are compared. Thus, the specific-heat data of isomorphic salts such as VN and ZrN and MgF_2 and ZnF_2 are not compatible with the corresponding-state approach although they are well fitted by the two-Debye method. It is apparent from Table 1 that the ratios of the two-Debye Θ's for such salts are not equal, and thus cannot be treated by corresponding-state considerations (see reference (12) for the mathematical details relating the two methods).

RESULTS

1. *Ferromagnetic*

Three ferromagnetic metals, Ni, Fe, and Gd, were analysed according to the methods given in the preceding discussion. We shall now discuss the separations and results for the three cases.

(a) *Separations.* For Ni, the electronic specific heat coefficient γ was determined by matching $\gamma T+bT^3$ between 15° and 30°K to the experimental specific-heat curve of BUSEY and GIAUQUE[13], and

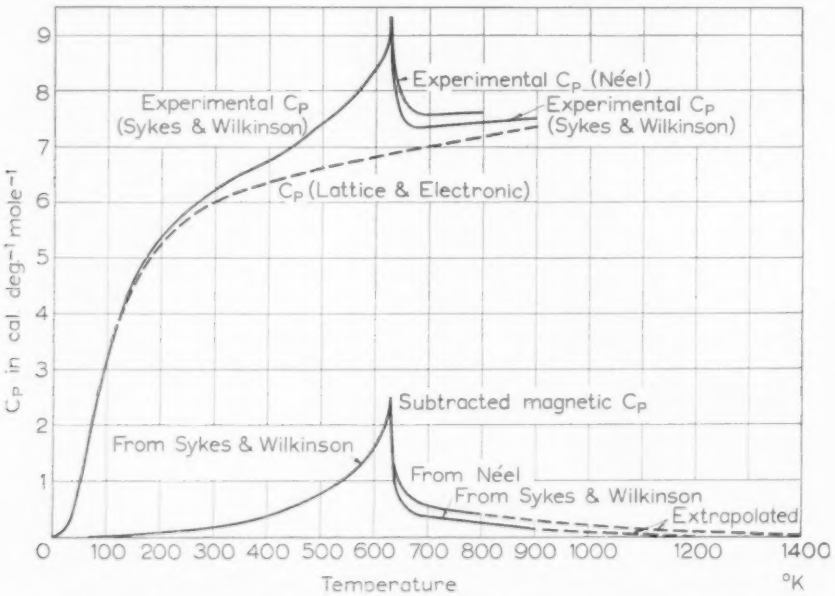

Experimental C_p (Néel)
Experimental C_p (Sykes & Wilkinson)
Experimental C_p (Sykes & Wilkinson)
C_p (Lattice & Electronic)
Subtracted magnetic C_p
From Sykes & Wilkinson
From Néel
From Sykes & Wilkinson
Extrapolated
C_p in cal. deg^{-1} mole^{-1}
Temperature
°K
0 100 200 300 400 500 600 700 800 900 1000 1200 1400
1 2 3 4 5 6 7 8 9

FIG. 5. Experimental specific heats of Ni after NÉEL, and SYKES and WILKINSON, together with the separated lattice and magnetic specific heats.

was found to be $12{\cdot}5\times10^{-4}$ cal. deg^{-1}mole^{-1}. The Debye Θ was then determined, as discussed in the section on ferromagnetic separations, by matching $\gamma T+(1+\alpha GT)f_D(\Theta/T)$ to BUSEY and GIAUQUE's data between 20° and 50°K, and was found to be 390°. Taking 0·6 effective electrons per atom then leads to a degeneracy temperature $T_0 = 4750$°K. In accordance with our previous discussion, the lattice specific heat is taken to be a Debye function with $\Theta = 390°$, corrected for lattice expansion as in eq. (8), and the electronic specific heat is given by eq. (6) with $\gamma = 12{\cdot}5\times10^{-4}$ and $T_0 = 4750$°K. The sum of the lattice and electronic specific heats is plotted in Fig. 5.

There is some disagreement in the literature on the experimentally determined total specific heat of Ni at high temperature, as indicated in Fig. 5. Two curves are given there for the specific heat above the Curie temperature, one due to SYKES and WILKINSON[14] and the other based on data of NÉEL[15]. Subtraction of the sum of the lattice- and electronic- specific-heat curves from the two total specific-heat curves leads to the two magnetic specific-heat curves in Fig. 5 plotted with solid lines. These two magnetic curves are extrapolated in a reasonable manner (dotted lines), and the areas W and total magnetic entropies of the resulting curves are calculated and listed in Table 3. An earlier measurement of the low-temperature specific heat of Ni by KEESOM and CLARK[16] leads to $\gamma = 17{\cdot}4\times10^{-4}$ and $\Theta = 413°$. If these values are used, together with a corresponding $T_0 = 3400°$, there results a lattice plus electronic curve

Table 3. Results of separation of magnetic specific-heat curves for ferromagnetic metals, together with parameters used in the separation. Values of J determined by application of spin-wave theory to low-temperature saturation magnetization measurements are also included.

Metal	Spin	T_c(°K)	W (cal/mole)	W/NkT_c	J (eV) (from W)	J (eV) (spin wave)	Exp. Entropy	Theor. Entropy	γ	Θ	Structure
Ni	0·30	630	420 ±50	0·57 ±0·07	0·010 ±0·0013	0·020	0·81 ±0·03	0·83	$12{\cdot}5\times10^{-4}$	390°	f.c.c.
Fe	1·11	1043	1930 ±120	0·93 ±0·06	0·011 ±0·0007	0·018	2·15 ±0·07	2·20	12×10^{-4}	432°	b.c.c.
Gd	3·56	292	806 ±40	1·38 ±0·07	0·00021 ±0·00001	0·00023	4·24 ±0·14	4·16	16×10^{-4}	152°	h.c.p.

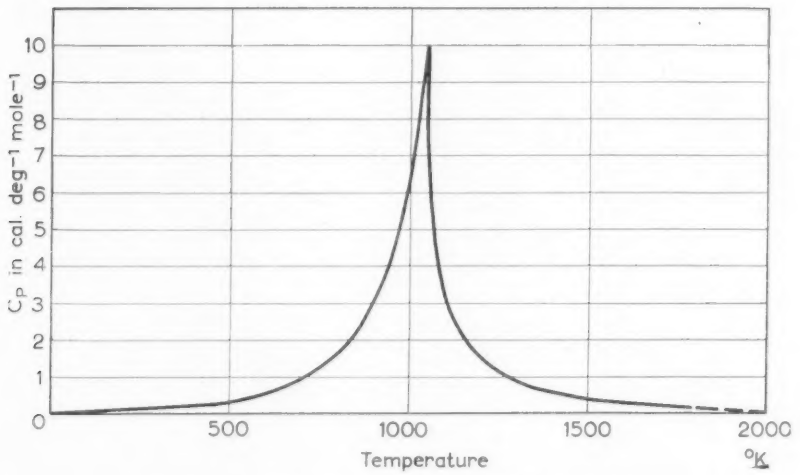


FIG. 6. The separated magnetic specific heat of Fe.

which crosses over the Sykes and Wilkinson high-temperature curve at about 900°K, and hence could not be correct if the Sykes and Wilkinson curve were the true curve. On the other hand, this lattice plus electronic curve taken with the Néel total curve leads to a reasonable magnetic curve which is not shown, but which lies quite close to the lower magnetic curve in Fig. 5, and hence has approximately the same area and entropy. It should be noted that while there is thus some question as to the best values of γ and Θ to use in the Ni subtraction, the corresponding uncertainty in our final values of W and magnetic entropy is not expected to be as large as the uncertainty due to the variations in the experimental high-temperature specific-heat data.

For Fe, KEESOM and KURRELMEYER[17] give the value of γ as 12×10^{-4}, obtained by matching the experimental specific heat between 1° and 4°K. Using our matching procedure, data of KELLEY[18] between 50° and 100°K leads to a Θ of 432°. The resulting degeneracy temperature T_0 is sufficiently high to make the second term in eq. (6) negligible. Subtraction of the electronic and lattice specific heats from the observed specific-heat curve[19] leads to the separated magnetic specific-heat curve shown in Fig. 6. The area W and the total magnetic entropy are listed in Table 3.

The experimental specific-heat data for Gd is due to GRIFFEL *et al*[20]. In the absence of low-temperature data, we have assumed[20] that γ for Gd is the same as that of lanthanum, namely 16×10^{-4}. Using our matching procedure, the experimental data between 15° and 25°K lead to a Θ of 152°. The validity of neglecting the magnetic contribution in this temperature range for Gd (with its relatively low Curie temperature) is discussed at the end of this section. The subtracted magnetic specific-heat curve is shown in Fig. 7, together with our high-temperature extrapolation. The resulting area W and entropy are listed in Table 3.

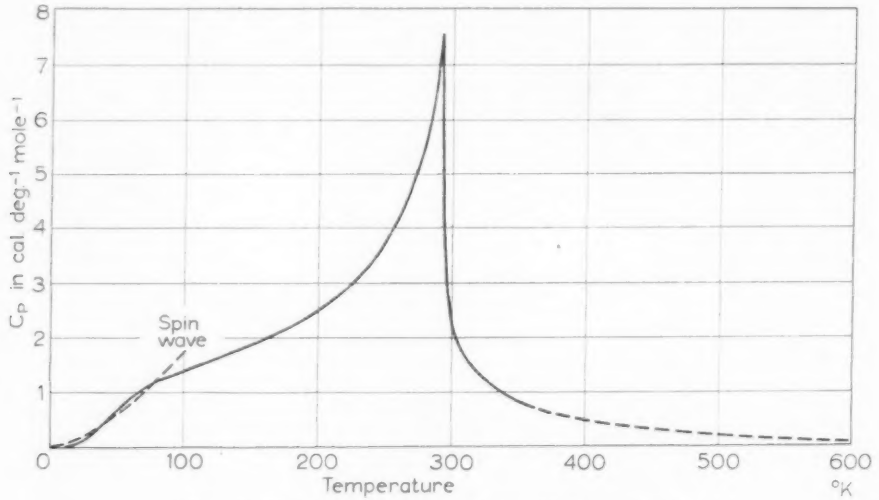


FIG. 7. The separated magnetic specific heat of Gd.

(b) *Results.* In the section on theory, it was tacitly assumed that in a given magnetic material all the spins were equal and of half-integral value. According to low-temperature saturation magnetization measurements, the spins of Ni, Fe, and Gd are 0·3, 1·1, and 3·56, respectively. In the case of Fe and Gd, the experimental spins are close enough to half-integral value to expect the discussion in the section on theory to apply if we take spin 1 for Fe and spin 7/2 for Gd. In the case of Ni, the spin is far from half-integral value, and Ni will be treated separately.

For half-integral spin s, the theoretical total entropy associated with a magnetic specific-heat curve is $R\ln(2s+1)$. The experimental entropies for Fe and Gd are compared in Table 3 with the theoretical entropies $R\ln 3$ and $R\ln 8$, respectively. It is seen that the experimental entropies are within 3 per cent of the theoretical values. As pointed out to us by Professor J. C. SLATER, this is evidence that the spin per atom is independent of temperature.

In order to compare experimental results for Ni with theory, we must first have some model to explain the non-half-integral spin of Ni, and then use this model to modify the results of the theoretical section. We have chosen the model for Ni discussed by VAN VLECK[21] where the $d^{9\cdot4}$ configuration of Ni is taken as 40 per cent $3d^{10}$ (spin 0) and 60 per cent $3d^{9}$ (spin 1/2) with the $3d^{9}$ and $3d^{10}$ configurations continually redistributing themselves among the lattice sites. For this model, the change in entropy in passing from an ordered magnetic state to a disordered magnetic state arises only from the $3d^{9}$ configuration, and is given by $0{\cdot}6R\ln 2$. It is seen in Table 3 that the experimental entropy for Ni is within 3 per cent of this value.

In keeping with our model for Ni we have made Weiss molecular-field estimates of the ratio W/kT_c for a magnetic material made up of a fraction β of spin s and a fraction $(1-\beta)$ of spin zero. We find that for this special case eq. (5) again results, where now N, the number of magnetic atoms, is β times the total number of atoms N_0, in the material. We therefore believe that the experimental value of W/NkT_c for Ni may be compared with the theoretical results, provided the effective spin is taken as 1/2 and N is taken as $0{\cdot}6N_0$.

Experimental values of the Curie temperature T_c are combined with the experimental values of

W to obtain experimental values of W/NkT_c. The values of W/NkT_c for the three ferromagnetic metals are listed in Table 3 and plotted in Fig. 1. The uncertainties indicated in the experimental W/NkT_c values in Table 3 and Fig. 1 follow from the estimated uncertainties in our determination of the magnetic energies of order. Fortunately, the area of a magnetic specific-heat curve, being an integrated quantity, is not too sensitive to small variations in the curve itself. The good agreement of the experimental entropies with the theoretically expected entropies is reassuring, although by no means an unambiguous indication of the accuracy of our experimental values of W, and hence of W/NkT_c.

The experimental values of W/NkT_c in Fig. 1 are to be compared with the theoretical calculations plotted there which are based on the Heisenberg exchange model and discussed in the section on theory. The results of the theoretical calculations of W/NkT_c as a function of spin are shown in Fig. 1 as continuous curves, although the calculations have actually been performed only for half-integral spins. (The cluster calculations were carried out only to spin 3. Beyond that, the theoretical curves are extrapolations.) In accordance with our previous discussion, the experimental values of W/NkT_c are plotted at spin 1/2 for Ni, spin 1 for Fe, and spin 7/2 for Gd. On this basis, it is seen that the group of theoretical curves in Fig. 1 is in general agreement with the experimental points, and that the cluster calculations appear to be an improvement over the simple Heisenberg theory.

Values of J, the exchange energy parameter, are determined for the three ferromagnetic metals from the experimental values of W through eq. (3) and are listed in Table 3. Another independent determination of J may be made from low-temperature saturation magnetization measurements. Spin-wave theory predicts[22] that the low-temperature saturation magnetization M_s is given by

$$M_s(T) = M_s(0)(1-CT^{3/2}) \qquad (13)$$

where $C = (0{\cdot}0587/\lambda s)(k/2Js)^{3/2}$ and λ is the number of lattice points per unit cell ($\lambda = 1$ for simple cubic, $\lambda = 2$ for body-centered cubic and hexagonal close packed, and $\lambda = 4$ for face-centered cubic). The values of J determined from the application of eq. (13) to the results of low-temperature saturation magnetization measurements are listed in Table 3[22, 23]. It is seen that for Fe and Ni the values of J determined from W are lower than those determined from the saturation magnetization measurements, but that for Gd there is good agreement.

We shall attempt to discuss the low-temperature region of the separated magnetic specific heats of Ni, Fe, and Gd in the light of the spin-wave prediction[24] that

$$C_v = \frac{0{\cdot}113R}{\lambda}\left(\frac{kT}{2Js}\right)^{3/2} \text{ cal/deg/mole of magnetic atoms.} \qquad (14)$$

For Ni, the separated magnetic specific heat is found to be consistent, on the basis of a log-log plot, with eq. (14) up to $T \sim T_c/2$ using $J = 0{\cdot}011 \pm 0{\cdot}003$ eV and $s = 1/2$, but not consistent with the assignment of $J = 0{\cdot}020$ eV, the low-temperature saturation magnetization value. In the case of Fe, the magnetic specific heat at these temperatures is too small a fraction of the total specific heat to allow an accurate match of eq. (14) to experiment. For Gd the calculated specific heat on the basis of spin waves, using $J = 0{\cdot}00021$ eV, is compared with our separated magnetic specific-heat curve in Fig. 7. It can be seen that the two curves are in approximate agreement up to $T \sim T_c/4$. Unlike Ni and Fe, the magnetic contribution to the specific heat of Gd in the range 15°–25°K is not negligible as originally assumed in our analysis. A redetermination of Θ based on the spin-wave magnetic contribution to the specific heat leads to a slightly higher Debye Θ of 158°. A re-evaluation of the magnetic energy and entropy on the basis of this higher Θ yields only a slight change in the values of these quantities given in Table 3. It might be noted that, while the experimental saturation magnetization[23] of Gd has a $T^{3/2}$ dependence up to $T \sim 0{\cdot}85T_c$, our magnetic specific-heat curve is inconsistent with a $T^{3/2}$ dependence above $T \sim T_c/4$.

2. *Antiferromagnetic*

Twelve antiferromagnetic salts were studied using subtraction procedures outlined previously. Our findings are given in Table 4. For convenience the salts are grouped according to spin. The parameters pertinent to our separation procedure (i.e., the Debye Θ's) are listed along with values of the

Table 4. Antiferromagnetic Compounds

Compound	Spin	T_c(°K)	W (cal)	W/NkT_c	JZ/k	Entropy	$R\ln(2s+1)$	θ's	Structure
*Cu Cl_2 $2H_2O$ [1]	1/2	4·3	7·54	0·87	15·1	1·31	1·39	No subt. nec.	
Ni F_2 [2]	1	73	139	0·95	69·4	2·19	2·19	306–700	Sn O_2
Ni Cl_2 [3]	1	53	94	0·89	47·1	2·10	2·19	325–390	Cd Cl_2
Ni O [4]	1	523	1016	0·97	507	2·30	2·19	425–900	Na Cl
†V Cl_3 [5]	1	30	206	0·94	28·2	2·05	2·19	350–425	As I_3
†V_2O_3 [6]	1	169	228	0·96	162	2·27	2·19	500–875	Cr_2 O_3
†MnO_2 [7]	3·2	90	184	1·06	42·5	2·60	2·77	475–950	Sn O_2
Fe F_2 [8]	2	79	193	1·22	24·1	3·18	3·22	240–625	Sn O_2
Fe O [9]	2	191	444	1·17	55·9	3·06	3.22	450–600	Na Cl
Mn F_2 [10]	5/2	67	173	1·29	13·8	3·54	3·58	220–575	Sn O_2
Mn O [9]	5/2	116	310	1·33	24·7	3·40	3·58	350–700	Na Cl
†$MnCl_2 4H_2O$ [11]	5/2	1·62	4·02	1·27	0·30	3·24	3·58	No subt. nec.	
Co O [12]	3/2	291		1·08	140				Na Cl
Fe Cl_2 [12]	2	24		1·20	7·2				Cd Cl_2
Mn Cl_2 [13]	5/2	2·0		1·30	0·42				Cd Cl_2

* High-temp. extrapolation.
† Low-temp. extrapolation.

(1) Friedberg S. A. *Physica* **18,** 714 (1952).
(2) Catalano E. and Stout J. W. *J. Chem. Phys.* **23,** 1284 (1955).
(3) Busey R. H. and Giauque W. F. *J. Amer. Chem. Soc.* **74,** 4443 (1952).
(4) Tomlinson J. R., Domas L., Hay R. G., and Montgomery C. W. *J. Amer. Chem. Soc.* **77,** 909 (1955).
(5) Shomate C. H. *J. Amer. Chem. Soc.* **69,** 220 (1947).
(6) Anderson C. T. *J. Amer. Chem. Soc.* **58,** 5641 (1936).
(7) Kellog K. K. and Moore G. E. *J. Amer. Chem. Soc.* **65,** 782 (1943).
(8) Calalano E. and Stout J. W. *J. Chem. Phys.* **23,** 1803 (1955).
(9) Todd S. S. and Bonnickson K. R. *J. Amer. Chem. Soc.* **73,** 3894 (1951).
(10) Stout T. W. and Adams H. E. *J. Amer. Chem. Soc.* **64,** 1535 (1942).
(11) Friedberg S. A. and Wasscher J. O. *Physica* **19,** 1072 (1953).
(12) Kittel C. *Introduction to Solid-State Physics* (J. Wiley & Sons, Inc., New York 1413/P189).
(13) Murray R. B. *Phys. Rev.* **100,** 1071 (1955).

energy of magnetic order W, the total magnetic entropy, and the exchange energy parameter J. Table 4 was compiled by using the specific-heat matching technique as opposed to the entropy-matching technique, which was not developed until after the bulk of our separations were performed. In the case of the fluorides, the two techniques were compared, and the results differed only by about 1 per cent in the values of W so obtained. We believe the entropy-matching technique is the preferable and simpler method whenever there is sufficient specific-heat data available both above and below the Curie temperature.

In order to estimate the error to be expected in our values of W when the specific-heat data covers the entire temperature range of interest (i.e., from a low-temperature value of at most about $T_c/4$ to a high-temperature value of at least $4T_c$), the following procedure was used. The experimental ZnF_2 entropy was matched at high temperatures with a two-Debye entropy function. The corresponding two-Debye specific-heat curve was then compared with the experimental ZnF_2 specific-heat curve (see Fig. 3). The difference in areas between the experimental and theoretical specific-heat curves was used to estimate the error in our technique of subtraction. If the difference in areas was taken with regard to sign, an error of about 1 per cent resulted over the temperature range of 20–200°K. If the absolute value of the difference in areas was taken, an error of 5 per cent was found. We therefore have taken the value of 5 per cent as an upper limit to the error in our entropy-matching technique on the assumption that the ZnF_2 specific-heat curve is representative of the lattice behavior of antiferromagnetic salts. We take the same estimate of error for our specific-heat matching technique.

In five of the twelve antiferromagnetic salts studied,

VOL. 1 1956/5

the experimental specific-heat data do not cover the entire range of interest. Four salts required some extrapolation on the low-temperature side of the Curie point and $CuCl_2 \cdot 2H_2O$ required extrapolation on the high-temperature side. The procedure in the cases of the four salts was to draw a smooth curve that goes through the lowest temperature specific-heat value and passes through $T = 0°K$ with zero slope. The salts requiring this type of extrapolation are marked by an asterisk next to the compound formula in Table 4. In view of the good entropy values obtained from these extrapolated curves by specific-heat matching, it is expected that a conservative estimate of the errors introduced in our final results by the low-temperature extrapolation procedure is about 10 per cent. The case of $CuCl_2 \cdot 2H_2O$ will be discussed in detail below.

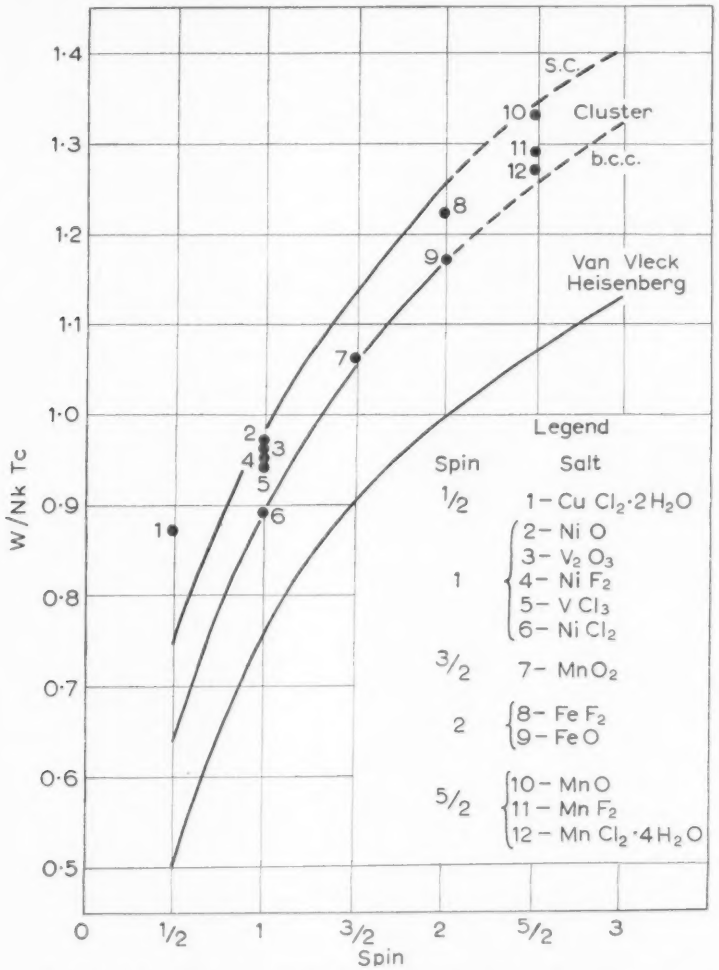


FIG. 8. Calculated ratios W/NkT_c as a function of spin for antiferromagnetic compounds based on the simple Heisenberg and the cluster theory. Values deduced from specific-heat measurements are included.

In Fig. 8 we plot the ratios W/NkT_c as a function of spin. It can be seen that there is good agreement between experiment and the cluster calculations of the Weiss-Li type. It should be noticed that the ratios W/NkT_c do not seem to depend on the details of the gross crystal structure. For comparison purposes the various crystal structures are listed in Table 4. It should be noted that for a

given salt the geometric arrangement of magnetic ions relative to one another is not necessarily the same as the gross crystal structure.

The case of $CuCl_2 \cdot 2H_2O$ is of especial interest. It is the only salt of spin 1/2 for which we have specific-heat data, and since the Curie temperature

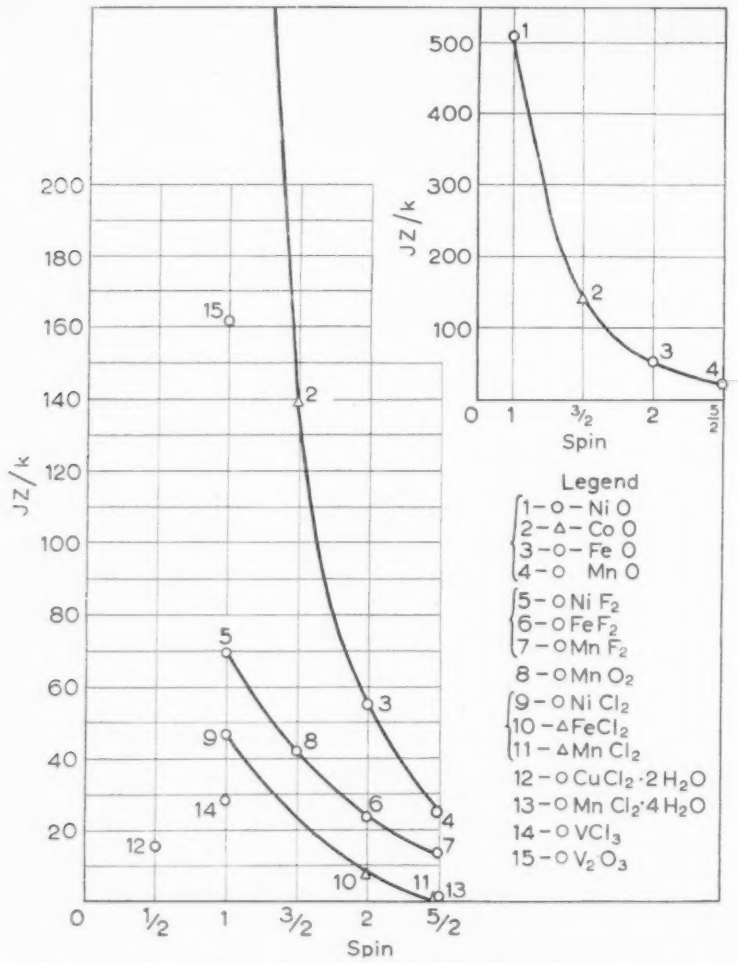

JZ/k
Spin
Legend
1 – o – Ni O
2 – Δ – Co O
3 – o – Fe O
4 – o Mn O
5 – o Ni F2
6 – o Fe F2
7 – o Mn F2
8 – o Mn O2
9 – o Ni Cl2
10 – Δ Fe Cl2
11 – Δ Mn Cl2
12 – o Cu Cl2 · 2 H2O
13 – o Mn Cl2 · 4 H2O
14 – o V Cl3
15 – o V2 O3

FIG. 9. JZ/k as a function of spin for antiferromagnetic compounds deduced from their specific heats. Solid curves connect isomorphic series.

occurs at about 4°K the lattice contribution may be neglected. This should minimize any possible ambiguity arising from our subtraction method. However, the experimental data unfortunately extend to a temperature only a little above the critical temperature and there is some ambiguity that arises from the process of extrapolating the data to higher temperatures. We have assumed that the data above the critical temperature have a $1/T^2$ dependence (see reference (a) in Table 4). Such an extrapolation produces a long tail in the specific-heat curve. Because we are at the high-temperature side of the specific-heat curve, the extrapolated curve adds proportionately less to the entropy than to the energy. In Table 4 it can be seen that the resulting entropy is about 6 per cent

less than the theoretical value. If the extrapolated curve were altered so as to produce enough entropy to agree with the theoretical value of $R \ln 2$, the energy value would lie considerably above the theoretical curves shown in Fig. 8. There are two possible explanations for an anomalously high value of W/NkT_c for spin 1/2 that come to mind. First, on the basis of spin-wave theory one expects a higher value of this ratio because of zero-point oscillations[25]. Second, W/NkT_c for a given spin increases as the number of nearest neighbors decreases (see Fig. 8: simple Heisenberg theory corresponds to an infinite number of nearest neighbors). Both effects are in the right direction to explain the anomaly, but further speculation is deferred until definitive data are available in the temperature region above the critical temperature.

Values of JZ/k based on our values of W and eq. (2) are plotted in Fig. 9. The most striking thing to notice about the behavior of JZ is that a smooth curve may be drawn through the various JZ values as a function of spin for isomorphic compounds. Thus, the chlorides,* fluorides, and oxides form three smooth curves of JZ vs. spin. For isomorphic compounds whose specific heats are not available but whose Curie temperature is known, it is possible to determine W and hence JZ by making use of an empirical curve drawn through the experimental W/NkT_c points in Fig. 8. Thus, JZ/k values for $FeCl_2$, CoO,† and $MnCl_2$ are calculated and lie on the smooth JZ/k curves of their respective isomorphic families as shown in Fig. 9. It should also be noted that MnO_2, which has the same structure as the fluorides‡ (see Table 4), seems to have a JZ value which lies on the same curve as the fluorides. This leads us to believe that isomorphic compounds with the same spin will have approximately the same Curie temperature. A program is now in progress to investigate this question for the various sulfides, selenides, and tellurides having the NiAs structure.

Acknowledgements—We wish to acknowledge several interesting discussions with Professor HARVEY BROOKS and Dr. R. K. NESBET. We thank Professor BROOKS in particular for his discussions pertaining to the problem of treating non-half-integral spin.

REFERENCES

1. HEISENBERG W. *Z. Physik* **49,** 619 (1928).
2. VAN VLECK J. H. *J. Chem. Phys.* **9,** 85 (1941).
3. VAN VLECK J. H. *Electric and Magnetic Susceptibilities* (Oxford University Press, Oxford, 1932) Chap. XII.
4. LIDIARD A. B. *Repts. Progr. Phys.* XVII, 201 (1954).
5. BROWN H. A. and LUTTINGER J. M. *Phys. Rev.* **100,** 685 (1955).
6. KRAMERS H. A. Leiden Supplement No. 83, 1936 (unpublished).
7. OPECHOWSKI W. *Physica* **4,** 181 (1937); **6,** 1112 (1938).
8. WEISS P. R. *Phys. Rev.* **74,** 1493 (1948).
9. LI Y. Y. *Phys. Rev.* **84,** 721 (1951).
10. WILSON A. H. *The Theory of Metals* (Cambridge University Press, 1953) Chapter VI.
11. STOUT J. W. and CATALANO E. *J. Chem. Phys.* **23,** 2013 (1955).
12. PASKIN A., HOFMANN J. A., TAUER K., and WEISS R. J. Submitted for publication in *J. Chem. Phys.*
13. BUSEY R. H. and GIAUQUE W. F. *J. Amer. Chem. Soc.* **74,** 3157 (1952).
14. SYKES C. and WILKINSON H. *Proc. Phys. Soc.* (London) **50,** 834 (1938).
15. NÉEL L. *Compt. Rend.* **207,** 1384 (1938).
16. KEESOM W. H. and CLARK C. W. *Physica* **2,** 513 (1935).
17. KEESOM W. H. and KURRELMEYER B. *Physica* **6,** 663 (1939).
18. KELLEY K. K. *J. Chem. Phys.* **11,** 16 (1943).

* J. W. LEECH and A. J. MANUEL (*Proc. Phys. Soc.* **69,** 220, 1956) have recently speculated that the transition metal chlorides may be ferromagnetic instead of antiferromagnetic. M. K. WILKINSON and J. W. CABLE (*Bull. Amer. Phys. Soc.* I, 1956, paper KA8) report that these chlorides consist of ferromagnetic layers antiferromagnetically arranged. This apparent disagreement does not affect our work. Our analysis of thermodynamic data and our conclusions in this section are independent of whether the salts are ferromagnetic or antiferromagnetic.

† The cobalt salts are of especial interest. While the calculated JZ value for CoO lies on the isomorphic oxide curve, those of CoF_2 and $CoCl_2$ do not lie on their respective isomorphic curves. It has already been pointed out by STOUT and CATALANO[11] that CoF_2 is anomalous in both its susceptibility and specific-heat behavior. Similarly we find $CoCl_2$ anomalous in its specific-heat behavior (TRAPEZNIKOWA *et al.*, *Phys. Z. Sowjet*, **9,** 237, 1936), and hence neither salt is represented in our work. However, it is of interest to note that while the JZ/k and W/NkT_c values of CoF_2 and $CoCl_2$ do not fit our curves on a spin 3/2 assignment, they do fit on a spin 1/2 assignment. As seen in Fig. 9, CoO fits the 3/2 spin assignment expected from free-ion considerations.

‡ It is not apparent how R. A. ERICKSON's preliminary observation (*Phys. Rev.* **85,** 745, 1952) that MnO_2 has a different magnetic structure from the fluorides relates to our observation on the JZ value of MnO_2.

19. DARKEN L. S. and SMITH R. P. *Ind. Eng. Chem.* **43,** 1815 (1951).
20. GRIFFEL M., SKOCHDOPOLE R. E., and SPEDDING F. H. *Phys. Rev.* **93,** 657 (1954).
21. VAN VLECK J. H. *Rev. Mod. Phys.* **25,** 221 (1952)
22. KITTEL C. *Introduction to Solid State Physics* (John Wiley & Sons, New York, 1953) 164.
23. ELLIOTT J. F., LEGVOLD S., and SPEDDING F. H. *Phys. Rev.* **91,** 28 (1953).
24. MOTT N. F. and JONES H. *The Theory of the Properties of Metals and Alloys* (Oxford University Press, London, 1936) 237. An algebraic error in their calculation has been corrected in eq. (14).
25. ANDERSON P. W. *Phys. Rev.* **83,** 1260 (1951).

VOL.
1
1956/5

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 61–64.

RESISTANCE AND PARAMAGNETISM AT THE SUPERCONDUCTING TRANSITION

JAMES C. THOMPSON*

The Rice Institute, Houston, Texas

(*Received* 9 *June* 1956)

Abstract—Experiments on the longitudinal magnetization and resistance of indium rods at the superconducting transition are described. Measurements were made under equilibrium conditions of large current, small magnetic field, and constant temperature. The correlation between resistance rise and occurrence of the "paramagnetic effect" is shown, and the form of the curve is found to be in accord with the theory proposed by H. MEISSNER. The maximum of the paramagnetism occurs when the resistance has approximately half its normal state value.

INTRODUCTION

IN the forty-five years since superconductivity was discovered, much information has been accumulated concerning the properties of the superconducting state of metals. Yet only within the last few years has a systematic study of the manner in which a metal enters the superconducting state been undertaken (FABER and PIPPARD[1]). The research to be reported here has dealt with the electric and magnetic properties of solid metal cylinders in the transition region between normal conduction and superconduction. In this mixed phase, usually called the intermediate state, the metal is neither wholly superconducting nor wholly normal conducting.

The established experimental facts are as follows. When the cylindrical specimen is below its critical temperature and a small longitudinal magnetic field is applied, the flux does not penetrate (Meissner effect). One says that nondissipative supercurrents flow on the surface in such a manner (in circles coaxial with the cylinder, in the case at hand) as to create a magnetic field which just cancels the external field, thereby maintaining zero induction in the specimen. If a small externally supplied current is passed down the axis of the specimen, no potential drop can be observed, since a superconductor has no resistance. We also know the externally sustained current is superficial, for otherwise there would be a nonzero magnetic induction within the cylinder. If the current is increased, the specimen eventually enters the intermediate state (Silsbee hypothesis). Under these conditions of current and field one finds a remarkable situation in the intermediate state. Magnetic flux exists inside the sample and a circular current component has been set up causing the longitudinal flux in the specimen to *exceed* that flux which would be expected from the external field alone. Previous work on this effect has been reported by STEINER[2], MEISSNER, SCHMEISSNER, and MEISSNER[3], THOMPSON and SQUIRE[4], MEISSNER and DOLL[5], MEISSNER[6], and THOMPSON[7]. This phenomenon has been called the "paramagnetic" effect in superconductors, and a theoretical description has been proposed by H. MEISSNER[6, 8]. As the current is increased above the critical current derived from the Silsbee hypothesis, the extra (paramagnetic) flux decreases. Finally, when the current far exceeds the critical current, the specimen has only its normal non-superconductive properties.

In the paramagnetic region (intermediate state), the effective conductivity is believed to be of a tensor character. The symmetry axis of the tensor is along a spiral path as a result of the mixture in the intermediate state of superconductive and normal metal (MEISSNER[8]). The current thus flows in a helical path and produces extra longitudinal flux much as a solenoid produces a longitudinal

* Shell Oil Company Fellow in Physics. Now a faculty member of the Department of Physics, University of Texas, Austin, Texas.

magnetic field. The maximum longitudinal flux occurs at the critical field corresponding to the temperature of the specimen. The total field is the vector sum (taken at the surface) of the azimuthal field (due to the current) and the applied longitudinal field.

We report here the simultaneous measurement of resistance characteristics and magnetic flux characteristics of a solid indium cylinder under the equilibrium conditions described above. W. MEISSNER[5] obtained similar results under dynamic conditions where temperature and current were continually varying. On the basis of the theory proposed by H. MEISSNER, and using the London[9] model for the distribution of normal and intermediate state material in the rod when the current slightly exceeds the critical current, we have derived expressions describing the resistance and flux behavior as functions of magnetic field, current, and temperature. Comparison of these expressions with the experimental results provides some verification of the proposed theory.

EXPERIMENTAL

The experimental apparatus has been adapted from that previously described (THOMPSON and SQUIRE[4], THOMPSON[7]). Fig. 1 shows schematically a solid cylinder of indium suspended in a liquid helium bath. A small homogeneous magnetic field can be applied parallel to the axis of the cylinder by means of a solenoid. A strong current (up to 20 amperes) can be sent down the axis of the specimen. The longitudinal flux content of the indium is compared with that in the superconducting lead rod (where $B = 0$) by means of the 4000-turn coil and its associated ballistic galvanometer. The deflection of the galvanometer is directly proportional to the longitudinal flux content of the indium. The ratio of the flux content of the indium (at a given field, current, and temperature) to the flux content of the indium when completely normal (at the same field and temperature) is called the effective relative permeability, and denoted $\tilde{K}$. $\tilde{K}$ is zero when the sample is superconducting; it is unity when the sample is normal; and it is greater than unity when the sample is in the intermediate state. The resistance of the indium was determined directly. The potential drop across the indium-lead sample and the current in it were measured. The lead remained superconducting throughout the experiment and thus had no resistance. A type K-2 Leeds and Northrup potentiometer was used for voltage determinations. Potentials were measured to within 1 per cent (relative to each other) near I_c; the absolute error was 3 per cent because of the intrinsic accuracy of the potentiometer on its lowest range.

The sample was a solid cylinder of 99·97 per cent pure indium, grown to be a single crystal. Its diameter was

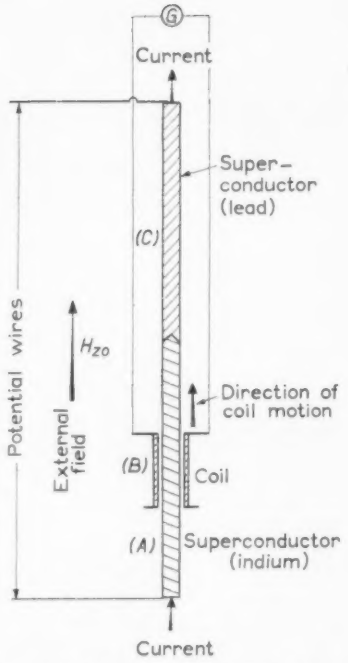


FIG. 1. Schematic diagram of sample and flux measuring system. The longitudinal flux content of the cylindrical specimen (A) was measured by quickly moving the concentric detection coil (B) from the indium to the lead superconductor (C) of zero flux content. The resulting deflection of the ballistic galvanometer (G) was directly proportional to the flux in the indium.

5·69 mm and its length 80 mm. The normal state resistance was $2{\cdot}4\times10^{-6}$ ohm yielding a residual conductivity of $1{\cdot}2\times10^{9}$ mho/meter. One infers from this low conductivity that the specimen was physically imperfect. Etching showed it to be composed of several crystals.

Measurements were made at constant temperature and external magnetic field. The current in the specimen was varied in small steps to take the specimen into and out of the superconducting state. A determination of flux content and resistance was made at each step under equilibrium conditions with constant H, I, and T.

A typical transition is shown in Fig. 2. The circles are experimental points and the solid curves are theoretical. Measurements taken where either of the other two variables (H, T) effected the transition would have yielded curves of essentially the same form.

Corroborative measurements were made using a slightly different technique on a different indium sample. A parallel path was provided for the current. This parallel circuit was composed of a superconducting tantalum wire wound in a helix, and also had a short copper strip placed in series so that the resistance of the parallel branch was finite (though small) and constant. As the

VOL. 1 1956/57

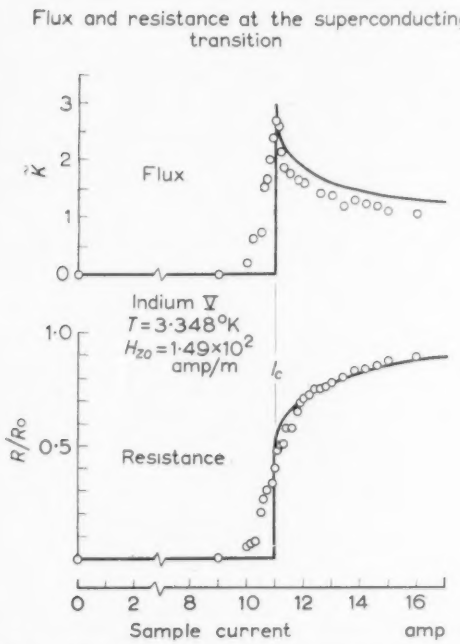


FIG. 2. Typical transition between normal and superconducting states. The upper portion of the diagram shows the relative permeability, and the lower portion shows the relative resistance. I_c is the critical current for this specimen.

resistance of the indium varied, the fraction of the current flowing in the tantalum varied. This fraction was detected and measured by measuring the resistance of a bismuth wire (diameter: 0·010 inch) inserted in the tantalum spiral. The magnetic field produced by the current in the tantalum increased the resistance of the bismuth (magnetoresistance). Transitions similar to Fig. 2 were obtained. The only difference was a steeper initial rise.

DISCUSSION

We have shown that an electric field exists in the metal when the metal is in the intermediate state. The resistance had approximately half of its normal state value when the paramagnetism was greatest (i.e., at $\tilde{K}_m$); no dependence on the magnitude of the maximum was observed. As the current increases from zero, the resistance begins to appear at about the same point that flux begins to appear in the specimen. Let us now look at the theory.

The theoretical treatment of the maximum of $\tilde{K}$, denoted $\tilde{K}_m$, has been given by H. MEISSNER[8], who also indicated the extension of the theory to currents above I_c[6]. For this latter derivation, one assumes that the cylindrical specimen is divided into two regions; an outer concentric sheath which is wholly in the normal state, and a central core which is in the intermediate state. The boundary is determined by the condition that the vector sum of external longitudinal field and the azimuthal field due to the current in the intermediate core be the critical field. Some resistance is restored to the core and as a consequence a portion of the current flows in the normal sheath. A more detailed calculation is given in the appendix. The result is: for the relative permeability

$$\tilde{K} = 1+(\tilde{K}_m-1)[2(I/I_c)^2-1-2(I/I_c)\sqrt{\{(I/I_c)^2-1\}}]$$

and for the resistance

$$R/R_0 = \tfrac{1}{2}[1+\sqrt{\{1-(I_c/I)^2\}}].$$

In the preceding equations, I is the current, I_c the critical current (i.e., the current at the peak of paramagnetism), R the resistance, and R_0 the normal resistance. When $\tilde{K}_m$ is determined from MEISSNER's theory, the solid curve in the upper portion of Fig. 2 is obtained; the resistance curve is in the lower portion of the figure.

The theoretical description of the flux behavior for currents greater than I_c has had mixed success. Fig. 2 shows one case where the general trend of $\tilde{K}$ is correct, though the experimental values lie below the theoretical curve. In some transitions $\tilde{K}$ compares well over the entire range of the current, and in others $\tilde{K}$ agrees with the theory at I_c yet lies well below the predicted value for all higher currents. So far as has been ascertained, there are no consistent differences in procedure or environment between the types of behavior cited. Previous work (THOMPSON[7]) has shown the theory to be incomplete for low currents or high external magnetic fields. The discrepancies in resistance and flux behavior are also believed due to inadequate theoretical treatment of $\tilde{K}_m$, not to experimental difficulties.

CONCLUSIONS

The presence of an electric field in the intermediate state has been demonstrated under equilibrium conditions, and this field is found to appear as the magnetic flux penetrates. The maximum of the paramagnetic permeability occurs when the resistance has approximately half of its normal state value.

The theory developed by H. MEISSNER is able to account for nearly all of the experimental observations. The use of the London model, when the current slightly exceeds the critical current, enables one to include the resistance behavior and the slow decrease of the paramagnetism in the theoretical framework.

Acknowledgement—The author would like to acknowledge the advice and assistance of Professors C. F. SQUIRE, W. V. HOUSTON, and H. E. RORSCHACH. The research program was supported in part by the National Science Foundation of the USA.

APPENDIX

As the current is increased above I_c, the extra flux does not disappear at once; rather it decreases slowly toward the normal state value. If one assumes, following London, that a sheath of normal metal is set up about a core of intermediate state material, then the decrease of extra flux can be understood. The radius of the intermediate-state core b is defined by

$$H_{z0}^2+(I_1/2\pi b)^2 = H_c^2, \tag{1}$$

where I_1 is the part of the total current, $I = I_1+I_2$, flowing in the core. I_2 is the current in the normal portion of the rod. From the Silsbee hypothesis, we have I_c given by

$$H_{z0}^2+(I_c/2\pi a)^2 = H_c^2 \tag{2}$$

where a is the radius of the sample. From MEISSNER's theory[6] equation (17), we have

$$b = I_c/(2\pi a\sigma_n E) \tag{3}$$

where E is the electric field and σ_n is the normal state conductivity. Combining (1) and (2) we find

$$I_1/2\pi b = I_c/2\pi a \quad \text{or,} \quad I_1/b = I_c/a. \tag{4}$$

Since the conductivity for a radius r, greater than b, is σ_n, I_2 will be given by

$$I_2 = 2\pi\int_b^a \sigma_n E r\,dr = \pi\sigma_n E(a^2-b^2). \tag{5}$$

We now equate the two expressions for I_1,

$$I_1 = I_c(b/a) = I-I_2 = I-\pi\sigma_n E(a^2-b^2) \tag{6}$$

and then eliminate $\sigma_n E$ by using (3). The result is a quadratic for b, which may be solved to yield

$$b = a[(I/I_c)\pm\sqrt{\{(I/I_c)^2-1\}}]. \tag{7}$$

Since $b \leqslant a$ for $I \geqslant I_c$, the negative sign is used.

Now let us define $\tilde{K}$ in a mathematical way. Since the external field is H_{z0}, the flux due to the external field alone is $\pi a^2\mu_0 H_{z0}$. In the case under discussion we must distinguish two regions. First, the normal region where only H_{z0} penetrates and the flux is thus $(a^2-b^2)\pi\mu_0 H_{z0}$. Second, the intermediate state region where $B \neq \mu_0 H_{z0}$, rather is given by the Meissner theory. The flux in this region will be given by

$$2\pi\int_0^b B_z(r)r\,dr. \tag{8}$$

Thus, since $\tilde{K}$ is defined as the actual longitudinal flux divided by the longitudinal flux due to the external field only, we find $\tilde{K}$ to be given by

$$\tilde{K} = \frac{\pi\mu_0 H_{z0}(a^2-b^2)}{\pi\mu_0 a^2 H_{z0}} + \frac{b^2}{a^2}\left[\frac{2\pi}{\pi b^2\mu_0 H_{z0}}\int_0^b B_z(r)r\,dr\right]. \tag{9}$$

The term in square brackets is called $\tilde{K}_m$ by MEISSNER[8], equation (1). If we now insert the value of b given by equation (7) the result is, as quoted earlier

$$\tilde{K} = 1+(\tilde{K}_m-1)\times \times[2(I/I_c)^2-1-2(I/I_c)\sqrt{\{(I/I_c)^2-1\}}]. \tag{10}$$

At this point a calculation of the resistance ratio is immediate. From (4) and (5), after b is eliminated by (3), we find

$$I = I_c^2/(4\pi a^2\sigma_n E)+\pi a^2\sigma_n E. \tag{11}$$

To simplify let $\Omega_0 = (\pi a^2\sigma_n)^{-1}$, which is the normal resistance per unit length, R_0/l. The solution of this quadratic in E yields

$$E = \frac{\Omega_0 I}{2}[1\pm\sqrt{\{1-(I_c/I)^2\}}]; I \geqslant I_c. \tag{12}$$

From (3),

$$E = \frac{I_c}{2\pi ab\sigma_n} \geqslant \frac{I_c}{2\pi a^2\sigma_n} = \frac{I_c\Omega_0}{2},$$

hence the positive sign is used before the radical in equation (12). The resistance may be defined by $R = (E/I)l$, and the final result is

$$R = (R_0/2)[1+\sqrt{\{1-(I_c/I)^2\}}]. \tag{13}$$

REFERENCES

1. FABER T. E. and PIPPARD A. B. *Progress in Low Temperature Physics* (New York, Interscience Pub., Inc.), Vol. I, (edited by C. J. GORTER), chapter IX (1955).
2. STEINER K. *Z. Naturforsch.* **4a**, 271 (1949).
3. MEISSNER, SCHMEISSNER and MEISSNER *Z. Phys.* **130,** 521, 529 (1951). *Z. Phys.* **132,** 529 (1952).
4. THOMPSON J. C. and SQUIRE C. F. *Phys. Rev.* **96,** 287 (1954).
5. MEISSNER W. and DOLL R. *Z. Phys.* **140,** 340 (1955).
6. MEISSNER H. *Phys. Rev.* **101,** 31, 1660 (1956).
7. THOMPSON J. C. *Phys. Rev.* **102,** 1004 (1956).
8. MEISSNER H. *Phys. Rev.* **97,** 1627 (1955).
9. LONDON F. *Superfluids* (New York, John Wiley and Sons), Vol. I, p. 120 (1950).

VOL. 1 1956/5

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 65–74.

ABSORPTION SPECTRA OF IMPURITIES IN SILICON—I
Group-III Acceptors

E. BURSTEIN, G. PICUS, B. HENVIS AND R. WALLIS

United States Naval Research Laboratory, Washington, D.C.

(*Received* 8 *June* 1956)

Abstract—The low-temperature absorption spectra of group-III impurities in silicon yield information about the ionization energies and excited states of the impurity centers. The optical values for the ionization energies of boron (0·046 eV) and indium (0·154 eV) are in good agreement with the thermal values, but the optical values for aluminum (0·67 eV) and gallium (0·71 eV) are appreciably larger than the thermal values. Variations in ionization energy among the group-III impurities are found to be accompanied by variations in the character of the excitation and photoionization spectra. The term schemes for the impurities are derived from the position of the excitation bands and the assumption that the $4p$ level is the same for all of the group-III impurities. Aside from the appearance of fine structure, the p-type states appear to fit a simple hydrogen model. From the positions of the p levels, an effective mass of 0·045 is calculated for holes using the expression for the energy levels of a hydrogen model. Variations in the oscillator strengths of the excitation bands among the group-III impurities are attributed to variations in the position and character of the $1s$ levels which arise from effects at the impurity atom that cause the simple effective-mass formalism to break down. The factors which should be included in a more complete theoretical treatment of the impurity centers are discussed.

INTRODUCTION

THE relatively low ionization energies of the group-III and group-V impurity elements in silicon and germanium were first explained in terms of a simple hydrogen model for the un-ionized impurity centers[1]. In this model the impurities are pictured as substitutional atoms with one more or one less nuclear charge to which the charge carriers are bound by a coulomb potential reduced by the dielectric constant of the medium. For bands with spherical constant energy surfaces, the ionization energy of a hydrogen-like impurity center is simply $E_I = (E_H/\epsilon^2)(m^*/m)$ where E_H is the ionization energy of hydrogen, m^* is the effective mass of the charge carrier, m is the mass of electrons in free space, and ϵ is the dielectric constant ($\epsilon = 11{\cdot}5$ for silicon and 16 for germanium). The hydrogen model for the impurity center is obtained by expressing the wave function of the bound charge carrier as a linear combination of Wannier functions[2]. For a medium with a high dielectric constant one obtains a differential equation, the effective-mass equation, which is essentially the Schrodinger equation for a hydrogen-like center with a potential given by $V = -e^2/\epsilon r$ and a scalar effective mass m^* in place of the electron mass in free space.

The early electrical data for silicon[3] indicated an ionization energy of 0·08 eV for boron acceptor centers and 0·06 eV for phosphorus donor centers, where as the early data for germanium[4] indicated essentially the same ionization energy of 0·01 eV for group-III acceptors and group-V donors. In silicon the difference between the ionization energy of boron, which was considered to be representative of other group-III elements, and that of phosphorus, which was considered to be representative of other group-V elements, was attributed, on the basis of the hydrogen model. to differences in the effective masses of electrons and holes. In germanium, on the other hand, the similarity in the ionization energies of group-III and group-V elements was attributed to similarities in the effective masses of electrons and holes. Estimates of the effective masses based on experimental values for the ionization energies and the assumption of a hydrogen

model indicated values of $0{\cdot}8m$ and $0{\cdot}6m$ for holes and electrons respectively in silicon and a value of $0{\cdot}2m$ for both electrons and holes in germanium.

Early infra-red studies of un-ionized impurity centers in silicon[5] appeared to confirm the simple hydrogen model for boron acceptor centers. At low temperatures, the un-ionized boron centers were found to exhibit three prominent excitation bands beyond the photoionization limit. The position of the bands, which were attributed to optical excitation of bound charge carriers from their ground state to higher energy (excited) states, were found to correspond closely to $1s$–$2p$, $1s$–$3p$ and $1s$–$4p$ transitions of a hydrogen-like center. The oscillator strengths of the excitation bands were however appreciably different from those of a hydrogen-like center. The optical data yielded a value of 0·046 eV for the ionization energy of boron centers, which was appreciably lower than the values of 0·08 eV reported by PEARSON and BARDEEN[3] from electrical measurements, and indicated an effective mass of $0{\cdot}45m$ for holes in silicon.

Recent electrical data now indicate that there are small but reproducible differences in the ionization energies among the group-III and among the group-V elements in germanium[6] and even larger differences among the group-III and group-V elements in silicon[7], which are not expected on the basis of an effective-mass treatment of the impurity centers (Table 1). Differences between the more recent values for the ionization energies of boron and phosphorus in silicon and the earlier values reported by PEARSON and BARDEEN[3] are attributed to peculiarities arising from nonuniformity or from unknown impurities in PEARSON and BARDEEN's samples[8].

Table 1

Impurity element	Thermal ionization energy of Group-III and Group-V impurities	
	Silicon (eV)	Germanium (eV)
Acceptors		
B	0·045	0·0104
Al	0·057	0·0102
Ga	0·065	0·0108
In	0·16	0·0112
Donors		
P	0·044	0·0120
As	0·049	0·0127
Sb	0·039	0·0096

Various attempts have been made to develop a more exact effective mass theory for the binding energy of the group-III and group-V impurity elements in silicon and germanium[9, 10, 11] by taking into account the detailed structure of the band edges which have been established by cyclotron resonance experiments. Due to the ellipsoidal character of the conduction band edges in silicon and germanium, the effective-mass theory for donors leads to a Schrodinger-type equation, with different masses in the three co-ordinate directions. The theory yields values for the ionization energies which are smaller than the experimental values and which are independent of the impurity atom involved. Similar deviations are also found between theoretical and experimental values of the hyperfine splitting of spin resonance lines for donors in silicon[12]. These discrepancies between theoretical and experimental values are attributed to the breakdown of the effective-mass formulation in the immediate vicinity of the impurity atoms.

An effective mass calculation has also been carried out for excited states of donor centers in silicon[11, 13, 14]. Since p-type states are not appreciably sensitive to details of the perturbing potential in the immediate vicinity of the impurity atom, the effective mass theory is presumed to hold quite well. The theoretical values for these energy levels are in good agreement with the values derived from optical data[15], although the latter still indicate small variations from impurity to impurity.

For acceptor centers, the effective-mass theory yields a set of six coupled differential equations which involve a mixing of wave functions from the different degenerate valence bands. A satisfactory variational solution has recently been given for germanium[16] in which the spin-orbit splitting at the top of the valence band is 30 times as large as the ionization energies of the acceptors so that the six coupled differential equations can be reduced to four as a good first approximation. Here again the calculated values for the binding energy are smaller than the experimental values, and the discrepancies are again attributed to the breakdown of the simple effective mass theory at the

VOL. 1 1956/5

impurity atom. A corresponding solution is not yet available for acceptor states in silicon where the analysis is complicated by the fact that the spin-orbit splitting is comparable in magnitude to the ionization energies of the acceptors.

Optical data now available for the group-III acceptors—boron, aluminum, gallium, and indium—in silicon show that the variation in the ionization energies among the group-III acceptors in silicon is accompanied by variations in the character of their excitation and photoionization absorption spectra[17]. The optical data yield fairly unambiguous values for the ionization energies of the impurities as well as information about the spectrum of excited states. Similar data have been obtained for group-V donor impurities and are described in the accompanying paper[18].

EXPERIMENTAL

The various silicon samples used in this investigation wer kindly supplied by J. A. BURTON, F. J. MORIN and R. G. SHULMAN of the Bell Telephone Laboratories. The impurity concentrations were estimated from the resistivities of the samples as well as from the room-temperature free-hole absorption at 2 microns.

The infra-red transmission measurements were carried out with a model 12-C Perkin-Elmer infra-red spectrometer equipped with interchangeable NaCl, KBr, and KRS-5 (TlBr-I) prisms. In order to reduce stray radiation a crystal quartz chopper was used in place of the opaque chopper to modulate the radiation at wavelengths longer than 5 microns. A further reduction of stray radiation was accomplished by using a LiF reststrahlen plate for the wavelength range of 15 to 25 microns and a CaF_2 reststrahlen plate for the wavelength range of 25 to 38 microns.

In order to carry out the transmission measurements at low temperatures where the carriers are bound to the impurity centers, the samples were cemented with colloidal silver paste over an aperture in the helium pot of a demountable, all metal, liquid-helium dewar equipped with KRS-5 windows[19]. The design of the dewar was such that the helium pot could be adjusted vertically to place either the sample or a reference aperture into the beam at a focal point in the entrance optics of the spectrometer. In order to prevent any appreciable warming of the sample by the incident radiation, the reststrahlen plates were placed between the source and the sample for measurements beyond 15 microns, and a germanium filter was placed in front of the dewar for measurements below 15 microns.

EXPERIMENTAL RESULTS

Characteristic low-temperature transmission spectra for silicon doped with boron, aluminum, gallium, and indium at concentrations of the order of $10^{16}/cm^3$ are given in Figs. 1, 2, 3 and 4, respectively. The positions in eV of the excitation bands and the photoionization absorption edges of the various impurities are recorded in Table 2 along with the values calculated for a hydrogen model.

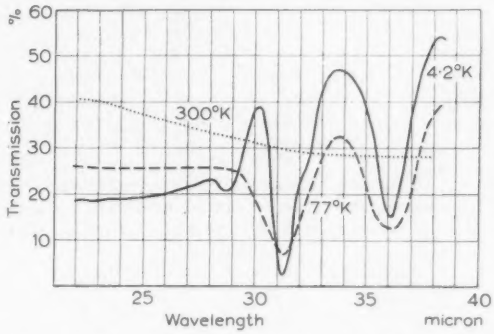


FIG. 1. Infra-red transmission spectra of boron-doped silicon at room temperature, liquid-nitrogen temperature and liquid-helium temperature.

$n_p = 5 \times 10^{15}/cm^3$. $t = 0{\cdot}53$ mm.

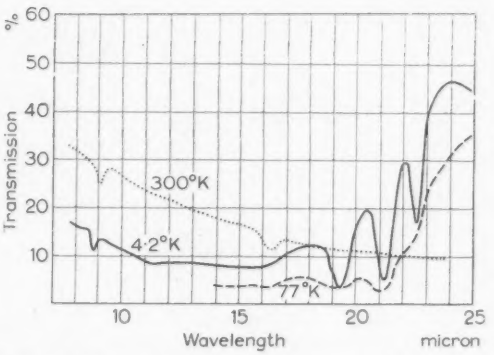


FIG. 2. Infra-red transmission spectra of aluminum-doped silicon at room temperature, liquid-nitrogen temperature and liquid-helium temperature. The bands at approximately 9 and 16 microns are lattice vibration bands.

$n_p = 6 \times 10^{16}/cm^3$. $t = 0{\cdot}35$ mm.

Boron, aluminum, and gallium each exhibit three prominent absorption bands at wavelengths longer than the photoionization limit. In indium only two such excitation bands are observed but it is presumed that a third one exists which is masked by the 8·8 micron lattice absorption band. The bands at 31 microns in aluminum- and gallium-doped silicon are due to the residual concentration

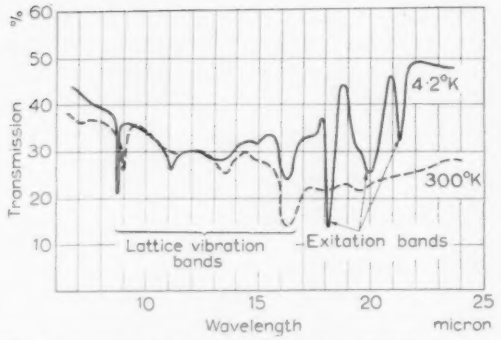


FIG. 3. Infra-red transmission spectra of gallium-doped silicon at room temperature and liquid-helium temperature. The bands from 8 to 19 microns are lattice vibration bands.
$n_p = 1{\cdot}0 \times 10^{16}/\text{cm}^3$. $t = 0{\cdot}89$ mm.

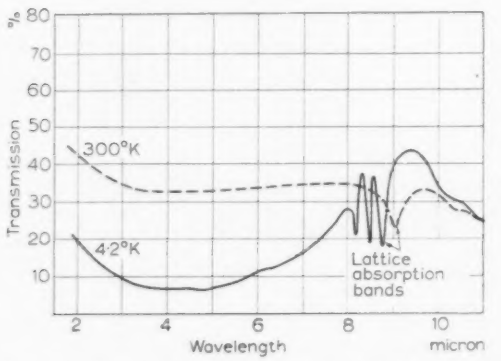


FIG. 4. Infra-red transmission spectra of indium-doped silicon. The bands beyond 8·5 microns are lattice vibration bands.
$n_p = 7 \times 10^{16}/\text{cm}^3$. $t = 1{\cdot}34$ mm.

Table 2

Impurity element	Position of bands in eV			
	1*s*–2*p*	1*s*–3*p*	1*s*–4*p*	1*s*–cont.* (E_1)
Hydrogen model	0·0345	0·0409	0·0431	0·0460
Boron	0·034	0·040†	0·043	0·046
Aluminum	0·055	0·059	0·064†	0·067
Gallium	0·058	0·06†	0·068†	0·071
Indium	0·141	0·146	0·151	0·154

* Calculated from position of 1*s*–4*p* band and assumption that 4*p* level is 0·003 eV below the bottom of the conduction band.

† Bands showing structure. Evidence for structure in the 1*s*–4*p* band of gallium centers is due to R. H. NEWMAN[20].

of boron present in the silicon. Lattice absorption bands also appear in the region from eight to 20 microns in all these samples.

As has been shown in an earlier work[5], the three excitation bands of boron centers correspond closely to 1*s*–2*p*, 1*s*–3*p*, and 1*s*–4*p* transitions of a hydrogen-like center and yield a value of 0·046 eV for the ionization energy of boron centers which is in good agreement with the thermal value of 0·045 eV (Table 3). Higher excitation bands do not appear since the bands have an appreciable width and merge with one another and with the photoionization continuum.

Table 3. Ionization Energies of Group-III Acceptors in Silicon

Impurity element	Thermal value* (eV)	Optical value (eV)
Boron	0·045	0·046
Aluminum	0·057	0·067
Gallium	0·065	0·071
Indium	0·16	0·154

* F. S. MORIN, S. P. MAITA, R. G. SHULMAN, and N. B. HANNAY [7].

The excitation bands of aluminum, gallium, and indium may similarly be identified with 1*s*–2*p*-, 1*s*–3*p*-, and 1*s*–4*p*-type transitions. It is the 1*s*–2*p* excitation band in indium which is apparently masked by the lattice absorption band. The positions of the excitation bands, unlike those for boron do not correspond to those of a hydrogen-like center. In the case of gallium, the 1*s*–3*p* band is actually composed of two overlapping bands.* This structure does not appear in the 1*s*–3*p* bands of aluminum and indium, but is found in the 1*s*–3*p* band of boron where it has been previously overlooked. NEWMAN[20] has independently noted the structure of the 1*s*–3*p* band of boron, and has actually been able to resolve it into two distinct

* When measurements were first carried out on gallium-doped silicon, it was first thought that the rather prominent shoulder of the 1*s*–3*p* band might be due to the superposition of the 19-micron lattice absorption band. However, subsequent measurements on a thinner, more highly doped sample showed no change in structure of the band while the lattice absorption bands had decreased appreciably in magnitude.

VOL. 1 1956/5

bands. A small shoulder is also present on the short wavelength side of the $1s$–$4p$ band of aluminum. NEWMAN also reports a small shoulder on the long wavelength side of the $1s$–$4p$ band of gallium. This structure is not resolved in our data.

Optical values for the ionization energies of aluminum, gallium, and indium can be estimated from the positions of their $1s$–$4p$ excitation bands if it is assumed that the $4p$ level, neglecting the existence of fine structure, is not appreciably influenced by the nature of the impurity center and therefore has the same position relative to the top of the valence band for all of the group-III impurities. Using the value of 0·003 eV derived from the boron data for the separation between the $4p$ level and the top of the valence band, we obtain a value of 0·067 eV for the ionization energy of aluminum, a value of 0·072 eV for gallium and a value of 0·154 eV for indium. The optical values for the ionization energies of aluminum and gallium are appreciably larger than the thermal values reported by MORIN and co-workers[7] (Table 3). The optical value for indium is however in good agreement with the thermal value.

Table 4

Impurity element	Oscillator strength of excitation bands*		
	$1s$–$2p$	$1s$–$3p$	$1s$–$4p$
Hydrogen model	0·416	0·079	0·029
Boron	0·04	0·08	0·04
Aluminum	0·01	0·04	0·09
Gallium	0·01	0·05	0·05
Indium	—	0·002	0·002

* Calculated from the integrated absorption cross sections, $\sigma(\nu_{ba}) = 1{\cdot}09 \times 10^{-16}(m/\epsilon^{1/2}m^*)f_{ba}$, using $m^* = 0{\cdot}45m$.

Approximate values for oscillator strengths of the excitation bands derived from the optical data are given in Table 4 which also includes the corresponding values for hydrogen-like centers. In view of the relatively large effective slit widths used in the measurements,* and the structure in some of the bands, the values quoted are only rough estimates. Nevertheless the data clearly indicate that the trend in the oscillator strengths of the excitation bands is markedly different from the theoretical values for a hydrogen-like center. For boron, aluminum, and gallium the oscillator

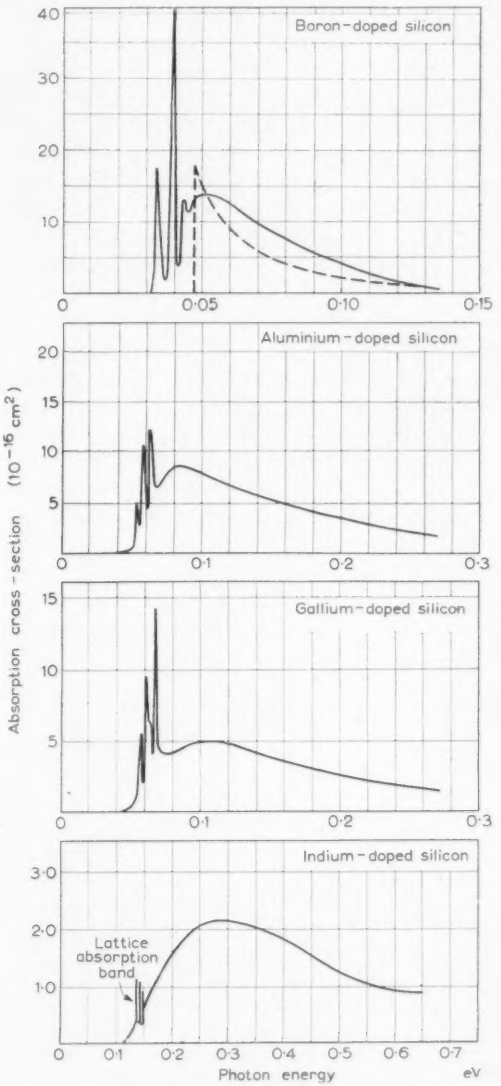


FIG. 5. Photoexcitation and photoionization absorption spectra of group-III impurities in silicon at liquid-helium temperature. For boron-doped silicon, the dashed line is the theoretical photoionization absorption spectrum for the corresponding hydrogen model.

* Effective slit widths are of the order of 0·15 microns in the region from 20 to 25 microns and increase from 0·5 up to 1·0 microns in the region from 25 to 38 microns.

strengths of the $1s$–$2p$ bands are an order of magnitude smaller than that for a hydrogen-like center, whereas the oscillator strengths for the higher bands are of the same order of magnitude. For indium, the oscillator strengths of the $1s$–$3p$ and $1s$–$4p$ bands are very much smaller than that of a hydrogen-like center and we may assume that this is also true for the $1s$–$2p$ band.

The photoionization absorption spectra also differ from that for a hydrogen-like center, the differences becoming more pronounced in going from boron to aluminum to gallium to indium (Fig. 5). Thus, the peak absorption in the photoionization spectrum does not occur at the ionization limit, but is shifted towards increasingly shorter wavelengths and the magnitude of the peak in the photoionization absorption spectra relative to the photoexcitation peaks increases in this sequence. On the other hand, the ratio of the magnitude of the $1s$–$4p$ excitation-band peak to the magnitude of the absorption at the ionization limit and, therefore the prominence of the $1s$–$4p$ band, shows the reverse trend.

The photoionization spectra of the group-III impurities in silicon show no evidence of the structure, which occurs in the corresponding spectra in germanium, involving transitions from the ground state to the V_3 valence band which is split away from the top of the band by spin-orbit coupling (Fig. 6). Since the splitting by spin-orbit

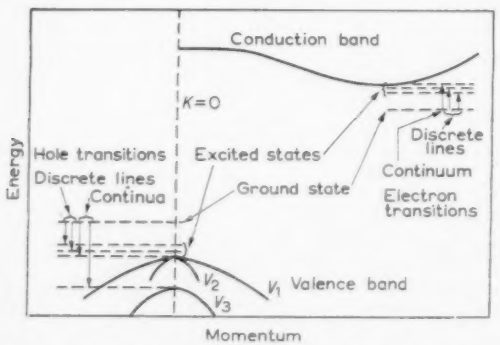


FIG. 6. Schematic energy-momentum diagram of bound states due to shallow donor and acceptor centers, showing the various optical transitions of bound carriers.

coupling in silicon is relatively small, the absorption limit for transitions to the V_3 band should lie close to that for transitions to the top of the valence band. The absence of structure is therefore presumed to be due to masking of the optical transitions to the V_3 band by those to the V_1 and V_2 bands.

The width of the excitation bands decreases appreciably in going from liquid-nitrogen temperature to liquid-helium temperature as can be seen in the data for boron and aluminum (Figs. 1 and 2) but there is no appreciable shift in the position of the bands. This indicates the absence of any appreciable temperature coefficient for the optical ionization energy at temperatures below that of liquid nitrogen. Analysis of the lattice vibration broadening of the energy levels[21] indicates that the optical ionization energies may be expected to be larger than the thermal ionization energies by about 0·001 eV because of Franck-Condon effects. This brings the optical and thermal values for boron into perfect agreement, a result which is undoubtedly fortuitous. The source of the large discrepancies between the optical and thermal values for the ionization energies of aluminum and gallium is not evident since details of the thermal measurements are as yet unpublished. In view of the unambiguous nature of the optical data the optical values for the ionization energies of aluminum and gallium are believed to be more reliable. A possible source of error in the thermal values for these impurities is the presence of uncompensated boron centers which would tend to give smaller slopes in the $\ln(np/T^{3/2})$ versus $1/T$ curves. Such uncompensated boron centers are present in sufficient concentration in the samples of silicon doped with aluminum and gallium used in this investigation to show up in the transmission spectra of these materials.

A term scheme for the group-III impurities derived from the optical data and the assumption of similar $4p$ levels for all of the impurities is given in Fig. 7. The term scheme for a hydrogen-like center having a similar $4p$ level (corresponding to a center with an ionization energy of 0·046 eV) is included for comparison purposes. We see that aside from the appearance of fine structure, the p-type levels as a whole correspond reasonably well to the p levels of a hydrogen-like center and that it is only the $1s$ state which is strongly influenced by the nature of the impurity atom. Higher s-type states may also be expected to be influenced by the nature of the impurity atom and should

VOL. 1 1956/5

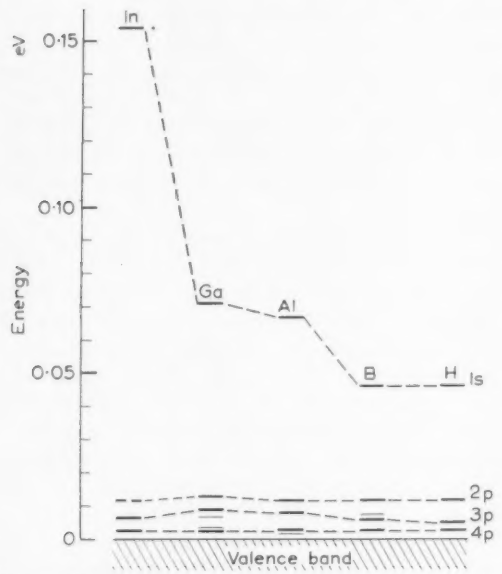


FIG. 7. Suggested term scheme for group-III impurities in silicon, showing levels whose positions are derived from the low-temperature absorption spectra.

exhibit variations among the impurities having the same trend as those of the 1*s* state. Small variations in the positions of the excited *p*-type levels do occur among the impurities and there appears to be a variation in the fine structure.* It is therefore likely that there is a variation in the positions of the 4*p* levels. Estimates of the optical ionization energies based on the position of the 4*p* level should, however, not be in error by more than 0·001 eV. The *p*-type states of the group-III impurities are apparently not appreciably influenced by the nature of the impurity atom and are therefore reasonably similar to one another period. Variations in the oscillator strengths among the impurities may accordingly be attributed predominantly to variations in the character of the 1*s* states.

Since the *p* levels appear to fit a hydrogen model reasonably well, an effective mass for holes can be calculated from their positions using the theoretical expression for the energy levels of a hydrogen-like center. This yields a value of 0·46*m* which does not differ much from the value of 0·5*m* for the average effective mass of holes in the V_1 band. It appears, therefore, that for the *p*-type states the V_1 band is the one which makes the major contribution.

NEWMAN[20] has recently investigated the effects of impurity concentration on the excitation spectra of boron, aluminum, and gallium in silicon. He finds that the spectra are independent of concentration below acceptor concentrations of $10^{16}/cm^3$. The excitation bands begin to broaden at concentrations above $10^{16}/cm^3$, and by concentrations of $10^{18}/cm^3$, the line structure is almost completely destroyed. The concentration broadening of the higher energy bands for a given impurity is found to occur at lower concentrations than for the lower bands. NEWMAN ascribes the concentration broadening of the bands to broadening of the excited levels by the overlapping of neighbouring impurity centers.

DISCUSSION

As shown by KITTEL and MITCHEL[9] and by LUTTINGER and KOHN[22] the bound acceptor states associated with group-III impurities in silicon and germanium are specified by a set of six coupled Schrodinger-like equations which arise from the three valence bands that are degenerate at $k = 0$ in the absence of spin-orbit interaction. In germanium, the spin-orbit interaction splits the V_3 valence band from the V_1 and V_2 valence bands by an energy which is large compared to the binding energy of the group-III acceptor, so that to a good approximation the set of six coupled effective-mass equations can be separated into a set of four coupled equations associated with the V_1 and V_2 bands and a set of two coupled equations associated with the V_3 band.

In silicon where the spin-orbit coupling is small all six coupled equations must be used to calculate the energies of the ground state and those excited states whose energies relative to the conduction band are comparable to the spin-orbit separation of the V_3 band from the V_1 and V_2 bands. For excited states whose energies relative to the conduction band are small compared to the spin-orbit separation it should be possible to calculate their properties with reasonable accuracy using the set of four coupled equations associated with the V_1 and V_2 bands. Without obtaining explicit solutions to the coupled effective-mass equations one can make some qualitative predictions concerning the

* Similar variations in the positions of the *p* states are also observed for group-V impurities in silicon where they are likewise due to effects at the impurity atom[18].

character of the effective-mass wave functions of the ground and excited states from symmetry considerations.

The coupled effective-mass equations are invariant with respect to the inversion $(x,y,z) \rightarrow (-x,-y,-z)$. The effective-mass wave function can then be classified as either even or odd with respect to inversion. The lack of rotational symmetry in the coupled effective-mass equations means that the wave functions for s-like states will have an admixture of d and possibly higher even atomic orbitals and that wave functions for p-like states will have an admixture of f- and possibly higher odd states. These considerations are borne out for the ground state by the calculations of KOHN and SCHECHTER[16] on the ionization energies of acceptor states in germanium which show that the ground state is predominantly s type with some d character. KOHN and SCHECHTER also point out that a reasonable estimate of the ionization energy for the ground state is given by a simple Bohr model with an effective mass derived from the mean curvature of the heavy hole band. It seems reasonable to expect that the energy level spectrum will not be drastically different from that for a hydrogen-like model and that in particular the sequence of the lower lying levels will be essentially that for hydrogen. The states can therefore be designated by the usual notation used for hydrogen; namely, $1s$, $2s$, $2p$, etc.

As an additional consequence of the lack of rotational symmetry the s and p states asociated with a given principal quantum number will not in general be degenerate. Furthermore the triple degeneracy of the p states themselves will be partially or wholly removed. Optical transitions are allowed between even states and odd states. The splitting of the p states means that the $1s$–$2p$, $1s$–$3p$ and $1s$–$4p$ transitions will each involve two or three frequencies rather than a single frequency. The presence of d character in the ground state will make possible an optical transition from the ground state to the states having predominantly f character. Under conditions where the $2s$ level is significantly populated, the $2s$–$2p$ transitions might be observed optically.

The optical absorption spectrum predicted by the effective-mass theory for degenerate bands is clearly more complicated than that predicted by a simple hydrogen-like model. Detailed calculations of the energy level spectrum for group-III impurities in silicon are being carried out by KOHN[23]. His preliminary calculations of the excited states do indicate the existence of a complicated absorption spectrum. However the theoretical spectrum has not yet been calculated with sufficient accuracy to allow a detailed comparison with experimental data.

The simple effective-mass theory in which the impurity potential is assumed to be simply of the form $-e^2/\epsilon r$ neglects effects at the central atom. Specifically, near the impurity atom the effective dielectric constant is different from the value characteristic of the bulk material and varies with the distance from the impurity site. Furthermore, the effective-mass formalism breaks down since the potential varies rapidly at the impurity atom. Recently BROOKS and FLETCHER[24] have calculated the wave function of the bound carriers by matching an effective mass type solution outside the impurity cell to a solution inside the impurity cell based on spectroscopic data for the impurity atom. For donors in silicon, they find an increased ionization energy over that predicted by the simple effective-mass theory; but their values are still not as large as those obtained experimentally. For acceptors in silicon the deviations from the hydrogenic model were also found to be small. For donors and acceptors in germanium their results do not differ from those of the simple effective-mass theory.

A complete theoretical treatment of impurity centers should also take into account the differences between the impurity-Si "covalent bond" and the Si–Si covalent bond including (1) the distortion of the lattice in the vicinity of the impurity atom which results from the difference in the relative strengths of these covalent bonds and from differences in the tetrahedral radii of the impurity and host lattice atoms, and (2) the polar character of the impurity-Si covalent bond which is related to differences in the electronegativity of the impurity and host lattice atoms. Along these lines SHULMAN[25] has carried out calculations of the energies of tight binding models for the impurity-silicon bond using Morse curves to approximate the interatomic potential. His calculations yield values for additional binding energies which increase sharply in going from gallium to indium in agreement with the observed trend in ionization energies.

VOL. 1 1956/57

To investigate effects arising from the specific nature of the impurity atom it may be desirable to treat the impurity center first as a many-electron problem. As a first approximation the ion-core electrons of the various atoms can be considered as coincident with the respective nuclei. The wave functions for the valence electrons can then be found by using, for example, the Hartree-Fock method. The wave functions thus found will show deviations near the impurity atom from the corresponding wave functions for the perfect crystal. These deviations may arise either from the displacement of the Si (or Ge) atoms near the impurity atom from perfect lattice sites or from the different effective nuclear charge of the impurity atom and the Si atoms. The distortion of the wave functions will lead to a difference in polar character in the impurity-Si and Si–Si bonds. One may reasonably expect the B–Si bond to be polarized with the negative charge at the Si atom—i.e., at the atom with the higher nuclear charge. The In–Si bond, on the other hand, may be expected to have the negative charge at the In atom. The polar character of the impurity-Si bond thus may be expected to have opposite effects in B and In.

The theory of acceptor states based on the hole concept generally assumes that the hole moves in the field of a potential having the form

$$V = V_p + U$$

where V_p is the potential of the perfect lattice and U is the potential that arises from the presence of the impurity. The considerations mentioned above indicate that U will not in general have the form $-e^2/\epsilon r$ commonly employed in effective-mass theory. Besides the effects due to the shift of the Si atoms near the impurity and to the polarity of the impurity-Si bond the potential of a hole is not expected to be $-e^2/\epsilon r$ within the ion-core of the impurity atom. These effects will lead to a variation of U from impurity to impurity.

If these features together with the rapid variation of U within or near the cell occupied by the impurity are incorporated in the theory, one may expect the states having predominantly s character to have their energies and wave functions rather strongly modified since such wave functions are relatively large at the impurity atom. The shift in the energy of the s states from impurity to impurity will have the further effect that the contribution of the V_3 band to the coupled effective-mass equations will vary in importance with the nature of the impurity.

The p states will probably be changed to a lesser extent than the s states since the former states have wave functions that are small at the center of the impurity atom. However, the shift of the Si atoms and the polarity of the impurity-Si bonds effectively increases the size of the impurity cell and will lead to a greater modification of the p states than would be expected from considerations of the ion-core potential alone.

The frequency of a given optical transition will in general vary from impurity to impurity because of the shift in energy levels. The oscillator strength may be expected to vary both because of the shift in energy levels and because of the change in the wave functions. The experimental data indicate that the increase in the depth of the ground state in going from boron to indium is correlated with a decrease in the oscillator strengths for transitions from the ground state to bound excited states. This correlation appears to be consistent with a considerably increased concentration of the 1s state at the impurity atom for indium which leads to an appreciably decreased overlap of the ground state wave functions with the p-state wave functions. It may be mentioned that the rather close agreement between the experimental data for boron and the predictions given by the simple hydrogen model probably arises from a cancellation of effects arising from deviations from the hydrogen model rather than from a lack of such deviations.

Acknowledgements—We wish to express our appreciation to P. EGLI and M. LAX for their active interest and encouragement throughout this investigation. We are indebted to J. A. BURTON, F. J. MORIN, and R. G. SHULMAN of the Bell Telephone Laboratories for profitable discussions and for their generous co-operation in exchanging information and in supplying the various silicon samples used in this investigation. We also wish to express our thanks both to W. KOHN and to R. NEWMAN for their friendly exchange of information on their respective work and for communicating their results to us prior to publication, and to N. SCLAR for stimulating discussions.

REFERENCES

1. BETHE H. *M.I.T. Radiation Laboratory Report No.* 43–12 (1942).
2. SLATER J. C. *Phys. Rev.* **36,** 1592 (1942).
3. PEARSON G. L. and BARDEEN J. *Phys. Rev.* **75,** 865 (1949).
4. See review paper by CONWELL E. *Proc. I.R.E.* **40,** 1327 (1952).
5. BURSTEIN E., BELL E. E., DAVISSON J. W., and LAX M. *J. Phys. Chem.* **57,** 849 (1953).
6. GEBALLE T. H. and MORIN F. J. *Phys. Rev.* **95,** 1085 (1954).
7. MORIN F. S., MAITA J. P., SHULMAN R. G., and HANNAY N. B. *Phys. Rev.* **96,** 833A (1954).
8. BURTON J. *Physica* **20,** 845 (1945).
9. KITTEL C. and MITCHEL A. H. *Phys. Rev.* **96,** 1488 (1954).
10. LAMPERT M. *Phys. Rev.* **97,** 352 (1955).
11. KOHN W. and LUTTINGER J. M. *Phys. Rev.* **98,** 915 (1955).
12. LUTTINGER J. M. and KOHN W. *Phys. Rev.* **97,** 883 (1955).
13. KLEINER W. H. *Phys. Rev.* **97,** 1722 (1955).
14. KOHN W. *Phys. Rev.* **98,** 1856 (1955).
15. PICUS G. S., BURSTEIN E., and HENVIS B. W. *Bull. Amer. Phys. Soc.* II, **1,** No. 3 (L3) 126 (1956).
16. KOHN W. and SCHECHTER D. *Phys. Rev.* **99,** 1903 (1955).
17. BURSTEIN E., PICUS G. S., HENVIS B. W., and LAX M. *Bull. Amer. Phys. Soc.* **30,** No. 2 (F10) 13 (1955).
18. PICUS G., BURSTEIN E., and HENVIS B. *J. Phys. Chem. Solids* **1,** 75 (1956).
19. DUERIG W. H. and MADOR I. *Rev. Sci. Instrum.* **23,** 421 (1952).
20. NEWMAN R. G. *Phys. Rev.* **103,** 103 (1956).
21. LAX M. and BURSTEIN E. *Phys. Rev.* **97,** 39 (1955).
22. LUTTINGER J. M. and KOHN W. *Phys. Rev.*, **97,** 869 (1955).
23. KOHN W. (private communication).
24. BROOKS H. and FLETCHER N. (unpublished) described in review paper by BROOKS H., Electrical properties of germanium and silicon, *Advances in Electronics and Electron Physics* **1,** 85 (1955).
25. SHULMAN R. G. *Phys. Rev.* **100,** 1260 (1955).

VOL.
1
1956/57

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 75–81.

ABSORPTION SPECTRA OF IMPURITIES IN SILICON—II
Group-V Donors

G. PICUS, E. BURSTEIN, AND B. HENVIS

United States Naval Research Laboratory, Washington, D.C.

(*Received* 8 *June* 1956)

Abstract—The infra-red absorption spectra due to neutral, *n*-type, impurity centers (P, As, Sb) in silicon have been measured at liquid helium temperature, and ionization energies and term schemes for these materials have been determined from the data. The ionization energies found are P: 0·0503 eV; As: 0·0533 eV; and Sb: 0·0426 eV. These values are approximately 10 per cent greater than the thermal ionization energies. The differences can be accounted for partly by (1) a Franck-Condon effect and (2) the presence of low-lying excited states. The positions of the excited *p* states of the impurity centers are all found to be roughly in the same positions relative to the conduction band, in good agreement with the results of the effective mass treatment of the impurity centers by KOHN and LUTTINGER. The variations observed between the various impurities are attributed to differences in the effective potentials in the immediate vicinity of the impurity ion cores. These differences are not taken into account in the effective mass theory.

INTRODUCTION

ALTHOUGH there is yet no adequate treatment of the energy levels of group-III acceptor centers in silicon, a fairly complete theoretical analysis has been carried out for the group-V donor centers.[1-7] The most definitive work is that of KOHN and LUTTINGER[6, 7] who determined the energies of the ground state and some of the excited states and approximate oscillator strengths for transitions between these states. They use the effective mass formalism to take into account the fact that the conduction band in silicon has six equivalent minima along (100) directions in k space and that the constant energy surfaces in the vicinity of these minima are ellipsoids with the major axes oriented along (100) directions. For the potential, KOHN and LUTTINGER use a coulomb potential reduced by the dielectric constant of the host silicon lattice ($V = -e^2/\epsilon r$). The term scheme they obtain is shown on the left in Fig. 1.

The notation used to characterize the levels is that of KOHN and LUTTINGER. The states are designated by the quantum numbers of the hydrogenic state that would result if the effective mass were isotropic. Since a bound state may be constructed in six different ways by using wave functions from each of the six conduction-band minima, each such state will be sixfold degenerate. The energies and the wave functions of the states are determined by a variational calculation which to date has been carried out only for the states

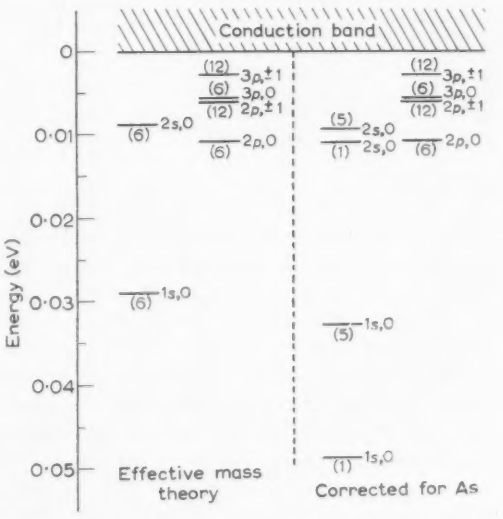


FIG. 1. Theoretical term scheme for an *n*-type impurity center in silicon calculated on the basis of the effective mass theory using a coulombic perturbing potential. Corrected term scheme for arsenic is on the right[6] [7].

tabulated in Fig. 1. The exact position of the s states for a particular impurity is not given by the use of a coulomb potential in the effective mass theory because of its neglect of effects at the impurity atom. From the difference between the theoretical and experimental values of the ionization energy, KOHN and LUTTINGER are able to make corrections in the positions of the s states for effects of the impurity atom, and the corrected term scheme for arsenic is shown at the right in Fig. 1. The degeneracy of the s states is seen to be split. The p-state wave functions have negligible amplitude at the impurity atom and so it is assumed that they will not be affected by the details of the potential in the immediate vicinity of the impurity. The energy separations of these excited states from the conduction band as found by KOHN and LUTTINGER are given in Table 1 together with the estimated relative intensities for transitions from the lowest $1s$ state.

Table 1. Separation of p states of an n-type impurity in silicon from the conduction band and relative intensity of transitions to the ground state

States	Separation from Conduction Band (eV)	Relative Intensity of Transition to Ground State
$2p, m = 0$	$-\cdot0109 \pm \cdot0002$	0·38
$2p, m = \pm 1$	$-\cdot0059 \pm \cdot0001$	1·0
$3p, m = 0$	$-\cdot0057 \pm \cdot0006$	0·04
$3p, m = \pm 1$	$-\cdot0029 \pm \cdot00005$	0·29

In this paper the experimentally observed infrared absorption spectra of phosphorus, arsenic, and antimony-doped silicon will be reported. Comparison of the data with the theory of KOHN and LUTTINGER enables us to construct term schemes for these materials and to estimate the ionization energies of the neutral impurity centers. Fairly good agreement is found between the experimental and theoretical results for the positions of the excited states although there are still variations from impurity to impurity. The optically determined ionization energies are in all cases larger than the thermal values and some speculations on the origin of the differences are presented.

RESULTS

The experimental procedures used are the same as those described in the preceding paper[8] on the p-type impurities. The samples used in this investigation were kindly furnished by J. A. BURTON and R. G. SHULMAN of the Bell Telephone Laboratories. Characteristic optical absorption spectra of phosphorus, arsenic, and antimony-doped silicon taken at liquid-helium temperatures in the wavelength region from 20 to 38 microns are shown in Fig. 2. The three curves shown are for samples which showed most prominently those features which appeared typical of the impurity present and independent of the history of the specimen. The features of the spectra on which attention will be focussed in this paper are the strong absorption lines at wavelengths of 26·2, 27·8, and 31·4 microns in phosphorus-doped silicon; 23·8, 24·5, 26·2, and 29·3 microns in arsenic-doped silicon; and 31·1 and 34·0 microns in antimony-doped silicon. The arsenic line at 31·2 microns, although moderately strong in the spectrum shown, did not vary from sample to sample with the impurity content as did the other four lines regarded as characteristic of the neutral arsenic center. Room-temperature free-electron concentrations for the materials used are listed in the figure. These values were obtained from measurements of the resistivity. Variations in the positioning of the helium pot in the dewar result in an uncertainty of about ± 3 per cent in the transmittance value plotted. Effective slit widths are of the order of 0·15 microns in the region from 20 to 25 microns and increase from 0·5 up to 1·0 microns in the region from 25 to 38 microns. In energy these correspond to $\Delta E \approx 0{\cdot}0035$ eV from 20 to 25 microns and $\Delta E \approx 0{\cdot}01$ eV in the KRS-5 prism region from 25 to 38 microns. These slit widths are too wide to permit more than a rough estimate of the half-widths of the absorption lines.

The observed absorption spectra are represented schematically as a function of energy (in eV) in Fig. 3. The energies corresponding to the various lines, as determined from the positions of the transmittance minima, are also given there. At liquid-helium temperature, practically all of the impurity centers will be in their ground state, and these absorption lines are the result of transitions from this state to higher excited states of the center. Differences in the energies of the absorption lines

VOL. 1 1956/5

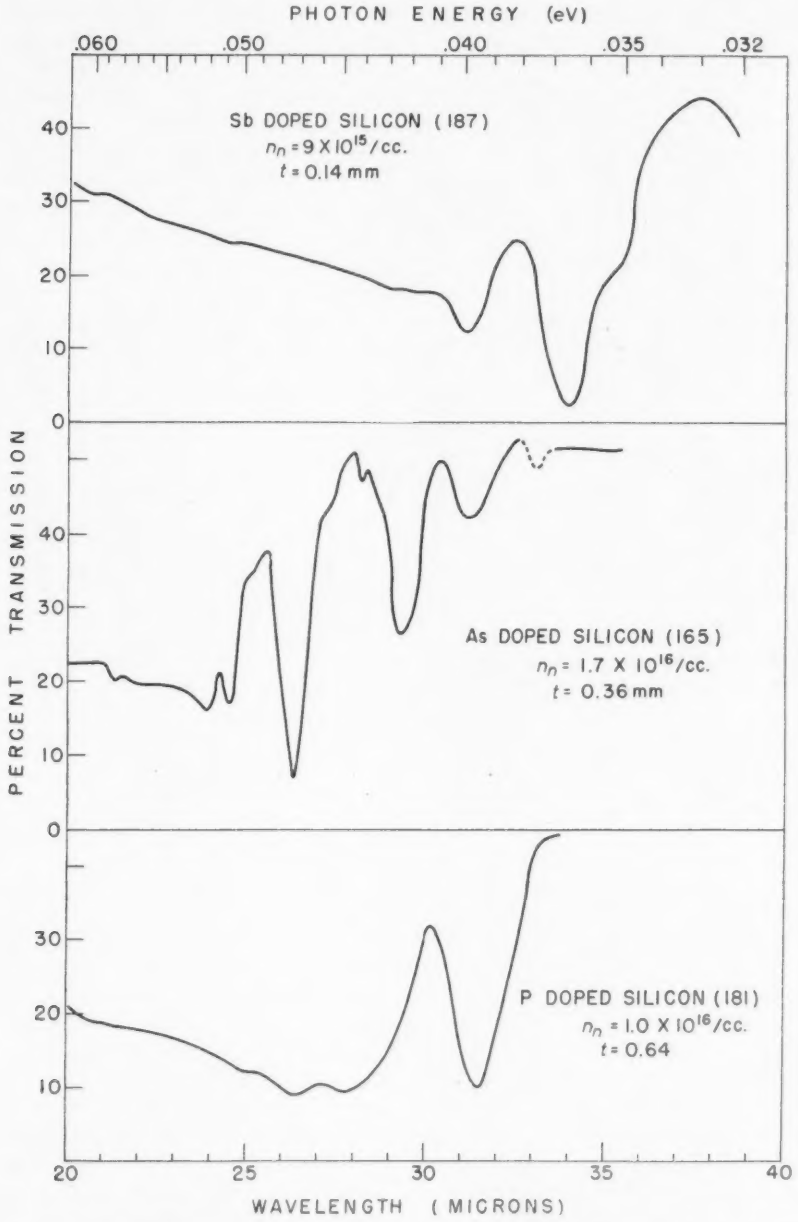


FIG. 2. Transmission spectra of *n*-type impurities in silicon as function of wavelength at liquid-helium temperature.

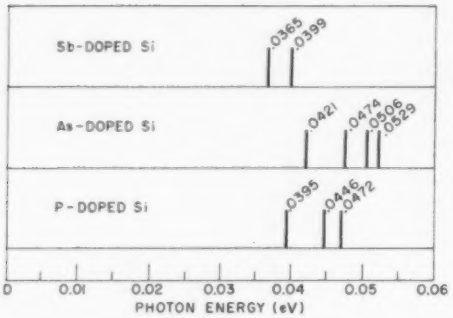


FIG. 3. Schematic representation of excitation absorption spectra of n-type impurities in silicon as a function of photon energy.

will therefore reflect the energy differences of the excited states. From the theoretical considerations of KOHN and LUTTINGER[6], it is reasonable to expect that a given excited p state has an energy that does not depend on which of the three doping elements is present. The energy separations of the excited states will then also be the same. In Table 2

Table 2. Comparison of theoretical spacing of excited states with observed spacing of absorption lines

Theoretical		*Experimental Spacings*		
States	Spacing eV	P	As	Sb
$2p,m = \pm 1 - 2p,m = 0$	0·0050	0·0051	0·0053	—
$3p,m = \pm 1 - 2p,m = \pm 1$	0·0030	0·0026	0·0032	0·0034
—	—	—	0·0015	—

the calculated spacings of the p states are compared with the spacings of the observed spectral lines for each of the doping elements. The experimental spacing, was placed in the same row with the theoretical spacing they most closely agreed with. From the excellent agreement that exists it is possible to identify the observed spectral lines with the theoretically determined states and thus to construct term schemes for the various impurity elements. These are presented in Fig. 4, where the calculated term scheme for an arsenic center is repeated for comparison. The exact position of the 1s ground state with respect to the bottom of the conduction band, i.e. the optical ionization energy, is determined by combining the experimental and theoretical data for the positions of the p states. The energy of a particular spectral line gives the separation of the corresponding p state from the ground state. By adding to this quantity the theoretical separation of the p state from the bottom of the conduction band we obtain a value for the ionization energy. The average of the several values thus found for each impurity

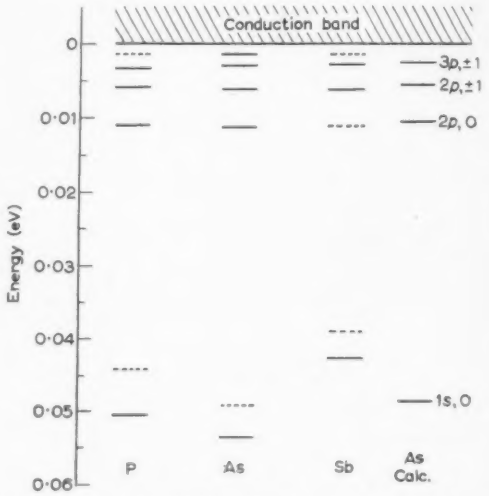


FIG. 4. Term schemes of n-type impurities in silicon. The calculated spectrum for As is shown on the right. The dashed lines just above the ground states represent the thermal ionization energies. The other dashed lines represent unobserved levels.

is given in Table 3 along with the thermal ionization energies determined by MORIN and coworkers[9]. The deviations listed for the optical ionization energies are the maximum differences found between the values determined from the various p states. The p levels in the term schemes are placed at distances above ground state corresponding to the energies determined from the optical absorption spectra. The uncertainty in the optical ionization energies given in Table 3 therefore reflects an uncertainty in the positioning of the bottom of the conduction band in the term schemes in Fig. 4. Only those levels between which transitions were actually observed are shown as solid lines in the term schemes. Thus the 2p,0 level

VOL. 1 1956/5

Table 3. Optical and thermal ionization energies of P-, As-, and Sb-doped silicon

Doping Element	Optical Ionization Energy (eV)	Thermal Ionization Energy (eV)	Per cent Diff.
P	0·0503 ± ·0006	0·044	12
As	0·0533 ± ·0006	0·049	8·1
Sb	0·0426 ± ·0006	0·039	8·4

in antimony-doped silicon is dashed because an absorption corresponding to the $1s$–$2p,0$ transition is not observed. On the other hand, there appears in arsenic-doped silicon a level above the highest calculated by KOHN and LUTTINGER which represents the upper level for the shortest wavelength transition observed in that material.

Referring back to Table 1, it can be seen that the theoretical $2p,\pm 1$ and $3p,0$ levels are separated by only 2×10^{-4} eV but that the former has a transition probability to the ground state about twenty-five times greater than the latter. From this we would expect to observe a small satellite on the short wavelength side of the line corresponding to the $1s$–$2p,\pm 1$ transition. There is some indication of the existence of such a satellite for the 26·3 micron line in arsenic-doped silicon, and the bulging shape of the 34·0 micron line in antimony-doped silicon may be due to the same cause. However, in both these instances there is even more prominent structure on the long wavelength side of the lines. This is particularly so in the antimony case. In the case of phosphorus-doped silicon the broadening is so severe that no structure can be observed in this particular line. The existence of a group of three lines at this position in the term scheme is explainable if it is assumed that the $m = \pm 1$ degeneracy of the $2p$ levels is split by some mechanism not included in the theory. Such a perturbation would have to account for the observed line intensities also.

In arsenic-doped silicon a transition (23·8 microns) is observed to a state higher than the $3p,1$ state, which is the highest for which theoretical information is available. It is most probable that this state is a $4p$ state but there exists a possibility that it may be the lowest $4f$ state. Transitions between such a state and the ground state would be allowed because the ground state has some d character. This transition is not observed in phosphorus- or antimony-doped silicon, but the short-wavelength lines in these materials are so broadened that it could be obscured very easily.

Because the absorption lines observed are quite broad and overlap appreciably it is difficult to make accurate estimates of relative intensities. In those few cases where the lines are sufficiently distinct, areas under the curves in the plots of absorption constant vs. energy were measured, and the results are presented in Table 4. The theoretical estimates of KOHN and LUTTINGER are presented there for comparison. In all cases the values are normalized by taking the relative probability for the $1s$–$2p,\pm 1$ transition as one. It is seen that to the limited accuracy possible there is considerable deviation from the calculated values. This is to be expected, to some extent, because the calculated transition probabilities will be more sensitive to the details of the wave functions used then will be the energies of the various levels. The wave functions are only approximate since they are not true eigensolutions, but the results of a variational calculation. Variations in the transition probabilities among the impurities no doubt result from differences in the perturbing potentials in the immediate vicinity of the impurity sites, a factor not taken into account in the effective mass theory. For a more detailed discussion of such effects see the preceding paper on p-type impurities[8].

Table 4. Comparison of theoretical and observed relative transition probabilities

Transition	Theoretical	P	As	Sb
$1s$–$2p,0$	0·38	0·5	0·3	—
$1s$–$2p,\pm1$	1	1	1	1
$1s$–$3p,0$	0·29	0·7	—	0·5

From Table 3 it can be seen that the thermal ionization energies are of the order of 10 per cent smaller than the optical ionization energies. There are two mechanisms which might help to account for the differences. One is a Franck-Condon effect which might result from differences in the lattice

configuration co-ordinates in the ground and excited states. If this situation existed there would also be an appreciable broadening of the observed lines. Although no accurate measurements can be made because of the large effective slit widths that had to be used it is evident from the absorption spectra that the sharpest lines were obtained for the arsenic-doped silicon and that the phosphorus-doped material gave the broadest lines. The broadening of the antimony-doped silicon absorption lines is intermediate. The percentage difference in the ionization energies is the smallest for arsenic, the impurity that gives the narrowest lines, and increases for the other two materials as does the broadening.

The presence of the higher "1*s*" state, predicted by theory to be 5-fold degenerate and to lie a small distance above the ground state, could also be responsible for introducing a difference between optical and thermal ionization energies. The effect of higher states of an impurity center on the temperature dependence of the carrier concentration has been investigated theoretically by SHIFRIN[10]. When such states are taken into consideration it is found that the slope of a plot of $\ln(n/T^{3/2})$ vs. $1/T$ is less than E_I, the energy separation of the ground state from the conduction band. The effect is negligible unless the higher states are of the order of kT above the ground state. For a separation of the 1*s* states of 0·016 eV (given by KOHN and LUTTINGER for As) the slope of a $\ln(n/T^{3/2})$ vs. $1/T$ curve for an impurity with the ionization energy of As would be 0·8 per cent lower than that expected. As the higher state approaches the ground state the slope of the curve would first decrease and then increase again to the value E_I when the states coalesced. The maximum change in slope would be only 5·5 per cent and would occur for a level separation of 0·005 eV; i.e., when the spacing of the levels is approximately equal to kT for the temperature range where the thermal measurements would be made (from 25°K to 50°K in the case of As- doped silicon[11]). The slight curvature that the presence of a low excited state would introduce into the logarithm of concentration vs. $1/T$ curves would be masked by a scatter in the experimental points of only 0·5 per cent. Detailed evaluation of the contributions of either of these mechanisms to the difference between the thermal and optical ionization energies depends on a knowledge of several quantities on which to date there is no information. In the first case the impurity center-lattice interaction has to be better understood and in the second case the exact position of the higher 1*s* state is needed. It can be seen, however, that both effects are in the proper direction.

CONCLUSIONS

The results reported in this paper and the preceding one[8] furnish further support for the usefulness and importance of optical techniques in the investigation of the electronic energy states in semiconductors. The close agreement between the experimentally determined term schemes for phosphorus, arsenic, and antimony centers in silicon and the theoretical energy levels calculated by KOHN and LUTTINGER on the basis of the effective mass formalism lends further support to the picture of the conduction bands utilized in the calculation and to the validity of the term scheme resulting from it. That is, the energies of the excited p states of an n-type impurity center from the fifth column of the periodic table as given by the theory are in accord with our present knowledge of the structure of the conduction band on the one hand and with the experimental data presented here on the other hand. However, the theory cannot account for the position of the ground state without correction and for the small variations in the positions of the p states that are observed experimentally. This points up the one major gap in the complete understanding of the nature of these impurity centers—there is little information about the form of the potential in the immediate vicinity of an impurity center.

Acknowledgements—We wish to acknowledge the active interest and encouragement of P. EGLI and M. LAX. We also wish to express our appreciation to J. A. BURTON, R. FLETCHER, F. J. MORIN, and R. G. SHULMAN for profitable discussion and for their generous co-operation in exchanging information and in supplying the various samples used in this investigation, to W. KOHN and J. M. LUTTINGER for their friendly exchange of information on their work, and to R. WALLIS for profitable discussions.

VOL.
1
1956/5

REFERENCES

1. KITTEL C. and MITCHEL A. H. *Phys. Rev.* **96,** 1488 (1954).
2. LAMPERT M. A. *Phys. Rev.* **97,** 352 (1954).
3. LUTTINGER J. M. and KOHN W. *Phys. Rev.* **96,** 902 (1954).
4. KOHN W. and LUTTINGER J. M. *Phys. Rev.* **97,** 883 (1954).
5. KLEINER W. H. *Phys. Rev.* **97,** 1722 (1954).
6. KOHN W. and LUTTINGER J. M. *Phys. Rev.* **98,** 915 (1955).
7. KOHN W. *Phys. Rev.* **98,** 1856 (1955).
8. BURSTEIN E., PICUS G., HENVIS B., and WALLIS R. *J. Phys. Chem. Solids* **1,** 65 (1956).
9. MORIN F. J., MAITA J. P., SHULMAN R. G., HANNAY N. B. *Phys. Rev.* **96,** 833A (1954).
10. SHIFRIN K. S. *Zhurna Tekhnicheskoi Fisiki* **14,** 43 (1944) (*J. of Tech. Phys. U.S.S.R.*)
11. MORIN F. J. and MAITA J. P. *Phys. Rev.* **96,** 28 (1954).

L.
/57

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 82–99.

ENERGY BAND STRUCTURE IN *P*-TYPE GERMANIUM AND SILICON

E. O. KANE

General Electric Research Laboratory, Schenectady, New York

(*Received* 5 *April* 1956)

Abstract—Energy-band calculations are made for the three valence bands in silicon and germanium in terms of the cyclotron resonance parameters. The energy in the band measured from $k = 0$ is not assumed small compared to the spin-orbit splitting so that parabolic bands do not result. The above calculation results from considering the first term of a perturbation expansion of the $k \cdot p$ and spin-orbit perturbations. The contributions from higher-order terms are examined and found to be important for germanium but not for silicon. Matrix elements for direct optical transitions between the valence bands are calculated from the cyclotron resonance constants. The free-carrier absorption is computed from the present band-structure calculations, and comparison is made with recent experimental data of R. NEWMAN for germanium. A correction to the split-off valence-band calculations is estimated, using the experimental data. Formulae are derived for degenerate perturbation theory with two perturbations of different orders acting.

VOL. 1 1956/57

1. INTRODUCTION

CONSIDERABLE progress has been made recently in understanding the band structure of silicon and germanium. This progress has resulted chiefly from cyclotron resonance measurements[1, 2, 3] and free-carrier absorption[4, 5] measurements. The theoretical discussion of degenerate bands near an energy extremum has been given by SHOCKLEY[6]. SHOCKLEY'S treatment has been extended to include the effects of spin-orbit coupling by DRESSELHAUS and others[1, 7, 8]. The analysis of free-carrier absorption in *p*-germanium by KAHN[9] has confirmed the interpretation of cyclotron resonance measurements and determined the spin-orbit splitting of the valence bands.

The general second-order perturbation equation for the energy bands including spin-orbit splitting has been given as a sixth-order determinant by DRESSELHAUS, KIP and KITTEL[1]. Explicit expressions for the energy bands have only been given for the case where the spin-orbit splitting is large compared to the energy in *k*-space measured from $k = 0$. KAHN'S[9] analysis also makes this approximation. In the present paper, the general secular equation is explicitly written as a cubic equation in Section 2. Numerical calculations are made for specific directions in silicon and germanium.

The sixth-order secular equation of reference one is derived as the leading term of a perturbation expansion of two perturbations of different orders. General expressions for the fourth-order terms are obtained in the appendix and are used in Section 3 to estimate the accuracy of the energies obtained from the leading term.

Matrix elements for direct optical transitions between the valence bands are calculated in Section 4 from the cyclotron resonance constants. The free carrier absorption is computed in Section 5 from the present band-structure calculations and comparison is made to recent experimental data of R. NEWMAN[10].

Many references will be made in this paper to the work of DRESSELHAUS, KIP and KITTEL[1]. Subsequently this paper will be referred to as DKK.

2. BAND STRUCTURE

The holes in the valence band of germanium and silicon are now known to possess wave functions which are in many respects similar to atomic *p* functions and may be thought to arise from such functions as we decrease the lattice constant of a fictitious germanium crystal from infinity to the actual value. The overlap of the individual atomic

p functions is, of course, large at the actual lattice constant and a very considerable modification of the atomic wave function naturally results so that the atomic approximation must not be taken too literally.

The holes of minimum energy lie at $k = 0$ in the Brillouin zone and in the absence of spin-orbit splitting would be six-fold degenerate, corresponding to the three p functions times the two spin functions, spin up and spin down. Spin-orbit splitting partially removes this degeneracy by lowering the two $j = 1/2$ bands with respect to the four $j = 3/2$ bands.

SHOCKLEY[6] derives the form of the energy surfaces in the neighborhood of $k = 0$ by a perturbation approach. In a periodic lattice the one electron wave functions can always be written as the well-known Bloch functions.

$$\psi_k = e^{ik\cdot r}u_k(r) \tag{1}$$

where $u_k(r)$ is cell periodic. The Schroedinger equation can then be written

$$[(p^2/2m)+V+(\hbar/m)k\cdot p]u_k(r) = \{E_k-(\hbar^2/2m)k^2\}u_k(r). \tag{2}$$

The term $(\hbar/m)k\cdot p$ is treated as a perturbation for determining u_k and E_k in the vicinity of $k = 0$ in terms of the complete set of cell periodic wave functions and energy eigenvalues at $k = 0$, which are assumed known. It is convenient to define an eigenvalue E'_k to simplify the form of eq. (2).

$$E'_k = E_k-(\hbar^2/2m)k^2. \tag{3}$$

In a lattice such as the diamond lattice where the inversion is a symmetry operation, the first-order matrix elements of the $k\cdot p$ perturbation vanish when the zero-order wave functions are taken at $k = 0$. If the band at $k = 0$ is nondegenerate, the surfaces of constant energy are spherical and the energy varies parabolically with the magnitude of k. If the band is degenerate at $k = 0$, the surfaces of constant energy may have more complicated shapes. The energy will still vary parabolically with the magnitude of k along a given direction in k space. SHOCKLEY[6] has analysed the form of the energy surfaces for cubic crystals using degenerate second-order perturbation theory. We repeat his derivation in more detail.

Consider a complete orthonormal set $|u_{J_i}\rangle$ of cell-periodic wave functions at $k = 0$ where J indexes different energies and i refers to degenerate functions for a given energy. The degenerate functions $|u_{J_i}\rangle$ for a given energy define a finite vector space where the vector c whose components are $c_1, c_2, \ldots c_n$ represents the function $[c_1u_{J_i}+c_2u_{J_2}+\ldots c_nu_{J_n}]$. Then the second-order energies and the correct zeroth-order wave functions are found by determining the eigenvalues and eigenvectors of the finite matrix equation

$$H_{kp}c = E'_kc \tag{4}$$

where the elements of H_{kp} are given by

$$\langle u_{J_i}|\mathcal{H}_{kp}|u_{J_k}\rangle = \sum_{L,n}{}' \frac{\langle u_{J_i}|(\hbar/m)k\cdot p|u_{L_n}\rangle\langle u_{L_n}|(\hbar/m)k\cdot p|u_{J_k}\rangle}{E_J-E_L} \tag{5}$$

The prime on the summation means that the term $L = J$ is excluded. The eigenvalue E'_k is related to the actual energy E_k by the definition (3).

The valence band wave functions at $k = 0$ in germanium and silicon transform according to the representation ϵ^+ in the notation of *DKK*. With respect to the basis functions $|\epsilon_1^+\rangle$, $|\epsilon_2^+\rangle$, $|\epsilon_3^+\rangle$ the Hamiltonian of eq. (4) may be written

$$H_{kp} = \begin{bmatrix} Lk_x^2+M(k_y^2+k_z^2) & Nk_xk_y & Nk_xk_z \\ Nk_xk_y & Lk_y^2+M(k_x^2+k_z^2) & Nk_yk_z \\ Nk_xk_z & Nk_yk_z & Lk_z^2+M(k_x^2+k_y^2) \end{bmatrix} \tag{6}$$

Symmetry arguments have been used to replace the sums over matrix elements in eq. (5) by the three constants L, M, and N. The quantities L, M, and N can be determined from cyclotron resonance measurements as discussed by DKK. Two sign ambiguities give four possible sets of L, M, N of which the two with positive N are easily discarded. In germanium the choice between the two remaining sets seems clearly to be the set selected by DKK. For silicon at the present time an ambiguity remains, although DKK state a preference.

The effects of spin-orbit coupling are most easily considered by regarding the spin orbit coupling interaction energy $\mathcal{H}_{so}$ as a perturbation.

$$\mathcal{H}_{so} = (\hbar/4m^2c^2)[\nabla V\times p]\cdot\sigma. \tag{7}$$

In the form of eq. (7) $\mathcal{H}_{so}$ is an operator on the Bloch function ψ_k. It can also be written in the following form as an operator on the cell periodic function u_k

$$\mathcal{H}'_{so} = (\hbar/4m^2c^2)[\nabla V\times p]\,\sigma + (\hbar^2/4m^2c^2)[\nabla V\times k]\cdot\sigma. \tag{8}$$

The first term is k independent and is analogous to the atomic spin-orbit splitting term. The second term is proportional to k and is the additional spin-orbit energy

coming from the crystal momentum. Rough estimates indicate that the effect of the second term on the energy bands is less than 1 per cent of the effect of the $(\boldsymbol{k}\cdot\boldsymbol{p})$ terms. The relatively greater importance of the first term comes essentially from the fact that the velocity of the electron in its atomic orbit is very much greater than the velocity of a wave packet made up of wave vectors in the neighbourhood of k. In subsequent discussion the k-dependent spin-orbit term will be neglected.

When spin-orbit effects are considered we take as our degenerate basis functions $|\epsilon_1{}^+\uparrow\rangle, |\epsilon_2{}^+\uparrow\rangle, |\epsilon_3{}^+\uparrow\rangle, |\epsilon_1{}^+\downarrow\rangle, |\epsilon_2{}^+\downarrow\rangle, |\epsilon_3{}^+\downarrow\rangle$; where $\uparrow$ and $\downarrow$ designate the spin functions spin-up and spin-down or $\binom{1}{0}$ and $\binom{0}{1}$. The Hamiltonian of eq. (4) then becomes a 6×6 matrix which may be written schematically

$$\begin{bmatrix} H_{kp} & 0 \\ 0 & H_{kp} \end{bmatrix}$$

where H_{kp} is the 3×3 matrix of eq. (6). In this representation the k-independent spin-orbit perturbation has the form

$$H_{so} = -\frac{\Delta}{3}\begin{bmatrix} 0 & i & 0 & 0 & 0 & -1 \\ -i & 0 & 0 & 0 & 0 & i \\ 0 & 0 & 0 & 1 & -i & 0 \\ 0 & 0 & 1 & 0 & -i & 0 \\ 0 & 0 & i & i & 0 & 0 \\ -1 & -i & 0 & 0 & 0 & 0 \end{bmatrix} \tag{9}$$

The symmetry properties of the matrix elements permit one to express all matrix elements in terms of a single constant Δ, the spin-orbit splitting. If we think of a fictitious germanium crystal in which we vary the lattice constant, the spin-orbit interaction Hamiltonian will only change through the multiplicative constant Δ. It is to be expected that in the tight-binding limit the spin-orbit Hamiltonian will be diagonalized by transforming to the Jm_j representation. Since the diagonalization is not affected by the value of Δ, the same transformation will diagonalize the spin-orbit interaction even though tight-binding is not a good approximation. The Jm_j transformation matrix is

$$U = \begin{bmatrix} -1/(2)^{\frac{1}{2}} & 0 & 0 & 0 & 1/(6)^{\frac{1}{2}} & 1/(3)^{\frac{1}{2}} \\ -i/(2)^{\frac{1}{2}} & 0 & 0 & 0 & -i/(6)^{\frac{1}{2}} & -i/(3)^{\frac{1}{2}} \\ 0 & (2/3)^{\frac{1}{2}} & -1/(3)^{\frac{1}{2}} & 0 & 0 & 0 \\ 0 & -1/(6)^{\frac{1}{2}} & -1/(3)^{\frac{1}{2}} & -1/(2)^{\frac{1}{2}} & 0 & 0 \\ 0 & -i/(6)^{\frac{1}{2}} & -i/(3)^{\frac{1}{2}} & i/(2)^{\frac{1}{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & (2/3)^{\frac{1}{2}} & -1/(3)^{\frac{1}{2}} \end{bmatrix} \tag{10}$$

The transformed matrix is

$$H'_{so} = U^{-1}H_{so}U$$

$$H'_{so} = \begin{bmatrix} \Delta/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & \Delta/3 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2\Delta/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Delta/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & \Delta/3 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2\Delta/3 \end{bmatrix} \tag{11}$$

The new basis functions φ_i are given in terms of the old basis function ϵ_j by the relation

$$\phi_i = \sum_{j=1}^{6} U_{ji}\epsilon_j. \tag{12}$$

The phases chosen for the Jm_j functions are not the conventional phases but have been selected to simplify the form of the Hamiltonian in the light of subsequent discussion. The ordering of wave functions implicit in the matrix (10) is

$$|j=3/2; m=3/2\rangle, |j=3/2; m=1/2\rangle,$$
$$|j=1/2; m=1/2\rangle, |j=3/2; m=-3/2\rangle,$$
$$|j=3/2; m=-1/2\rangle, |j=1/2; m=-1/2\rangle.$$

A theorem due to KRAMERS states that for any odd electron system in the absence of an external magnetic field all levels are necessarily doubly degenerate. The proof for a one-electron system consists in showing that the operator

$$K = -i\sigma_y\mathcal{C} \tag{13}$$

commutes with the Hamiltonian including spin-orbit interactions and has the property that $K^2 = -1$. $\mathcal{C}$ is the complex conjugation operator, σ_y is the spin operator $\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$. The assumption $K\psi = a\psi$ then leads to $K^2\psi = a^*a\psi = -\psi$ which is a contradiction. Therefore every state is at least doubly degenerate. When KRAMERS' operator is applied to the Bloch functions of a crystal it leads to the results that wave functions with k and $-k$ are degenerate even when the crystal does not possess inversion symmetry. If the crystal does possess inversion symmetry we consider the operator

$$\mathcal{K} = -i\sigma_y\mathcal{C}J \tag{14}$$

where J is the inversion. (For the diamond lattice J is to be replaced by $(\tau|J)$ where τ is the nonprimitive translation (1/4, 1/4, 1/4). This operator also commutes with the Hamiltonian and has the property $\mathcal{K}^2 = -1$. Furthermore the $\mathcal{K}$ operator leaves $e^{ik\cdot r}$ invariant so that in this case the bands are doubly degenerate. It is

VOL. 1 1956/57

easily seen that $\psi = \begin{bmatrix} W_a \\ W_b \end{bmatrix}$ and $K\psi = \begin{bmatrix} W_b^* \\ -W_a^* \end{bmatrix}$ are always orthogonal, also $\mathcal{K}\psi$ is normalized if ψ is normalized. If ψ_2 is chosen orthogonal to both ψ_1 and $\mathcal{K}\psi_1$, it is easily seen that $\mathcal{K}\psi_2$ will also be orthogonal to both ψ_1 and $\mathcal{K}\psi_1$. Hence it is always possible to choose a basis of the form $|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle, |\mathcal{K}\psi_1\rangle, |\mathcal{K}\psi_2\rangle, |\mathcal{K}\psi_3\rangle$. We note the following identities

$$\langle \mathcal{K}\psi_i | \mathcal{K}\psi_j \rangle = \langle \psi_i | \psi_j \rangle^* \tag{15}$$

$$\langle \mathcal{K}\psi_i | \mathcal{H} | \mathcal{K}\psi_j \rangle = \langle \psi_i | \mathcal{H} | \psi_j \rangle^* \tag{16}$$

$$\langle \psi_i | \mathcal{H} | \mathcal{K}\psi_j \rangle = -\langle \psi_j | \mathcal{H} | \mathcal{K}\psi_i \rangle. \tag{17}$$

As a consequence of these identities the Hamiltonian matrix will take the following form when referred to a Kramers' basis

$$H = \begin{bmatrix} G & \Gamma \\ -\Gamma^* & G^* \end{bmatrix} \tag{18}$$

where Γ has the property $\Gamma^\dagger = -\Gamma$. In the present 6-dimensional case Γ has the form

$$\Gamma = \begin{bmatrix} 0 & \gamma_1 & \gamma_2 \\ -\gamma_1 & 0 & \gamma_3 \\ -\gamma_2 & -\gamma_3 & 0 \end{bmatrix} \tag{19}$$

The unitary transformation matrix from one Kramers' basis to another has the form

$$U = \begin{bmatrix} S & R \\ -R^* & S^* \end{bmatrix} \tag{20}$$

where $H' = U^{-1}HU$. We will use the standard forms (18) and (20) in order to write 6×6 matrices as two 3×3 matrices. The transformation matrix (10) has the standard form (20). The $\pm m$ values for a given j are Kramers degenerate. This result follows from the fact that Kramers' operator $-i\sigma_y\mathcal{C}$ is a time reversal operator.

We write the 6×6 $k\cdot p$ Hamiltonian corresponding to eq. (6) in the Jm_j representation. The total Hamiltonian is then the sum of the $k\cdot p$ Hamiltonian and the spin-orbit Hamiltonian of eq. (11). The zero of energy is taken at the top of the valence band so that $\Delta/3$ must be subtracted from the spin-orbit Hamiltonian. The total Hamiltonian is

$$G = \begin{bmatrix} \frac{H_{11}+H_{22}}{2} & \frac{-H_{13}+iH_{23}}{(3)^{\frac{1}{2}}} & \frac{H_{13}-iH_{23}}{(6)^{\frac{1}{2}}} \\ \frac{-H_{13}-iH_{23}}{(3)^{\frac{1}{2}}} & \frac{H_{11}+H_{22}+4H_{33}}{6} & \frac{H_{11}+H_{22}-2H_{33}}{3(2)^{\frac{1}{2}}} \\ \frac{H_{13}+iH_{23}}{(6)^{\frac{1}{2}}} & \frac{H_{11}+H_{22}-2H_{33}}{3(2)^{\frac{1}{2}}} & \frac{H_{11}+H_{22}+H_{33}}{3} - \Delta \end{bmatrix} \tag{21}$$

$$\Gamma = \begin{bmatrix} 0 & \frac{-H_{11}+H_{22}+2iH_{12}}{2(3)^{\frac{1}{2}}} & \frac{-H_{11}+H_{22}+2iH_{12}}{(6)^{\frac{1}{2}}} \\ \frac{H_{11}-H_{22}-2iH_{12}}{2(3)^{\frac{1}{2}}} & 0 & \frac{H_{13}-iH_{23}}{(2)^{\frac{1}{2}}} \\ \frac{H_{11}-H_{22}-2iH_{12}}{(6)^{\frac{1}{2}}} & \frac{-H_{13}+iH_{23}}{(2)^{\frac{1}{2}}} & 0 \end{bmatrix} \tag{22}$$

The symbols H_{ij} refer to the elements of the matrix given in eq. (6).
We make the definitions

$$Xe^{i\chi} = -H_{11}+H_{22}+2iH_{12} \tag{23}$$

$$Ye^{i\eta} = H_{13}-iH_{23} \tag{24}$$

$$Z = H_{11}+H_{22}-2H_{33}. \tag{25}$$

The number of nonzero terms in Γ can be reduced by making the following unitary transformation

$$S_1 = \frac{1}{(2)^{\frac{1}{2}}} \begin{bmatrix} \exp i\left[\frac{\chi+\eta}{2}-\frac{\pi}{4}\right] & 0 & 0 \\ 0 & \exp i\left[\frac{\chi-\eta}{2}-\frac{\pi}{4}\right] & 0 \\ 0 & 0 & \exp i\left[\frac{\chi-\eta}{2}-\frac{\pi}{4}\right] \end{bmatrix} \tag{26}$$

$$R_1 = \frac{1}{(2)^{\frac{1}{2}}} \begin{bmatrix} -\exp i\left[\frac{\chi+\eta}{2}-\frac{\pi}{4}\right] & 0 & 0 \\ 0 & -\exp i\left[\frac{\chi-\eta}{2}-\frac{\pi}{4}\right] & 0 \\ 0 & 0 & -\exp i\left[\frac{\chi-\eta}{2}-\frac{\pi}{4}\right] \end{bmatrix} \tag{27}$$

The transformed Hamiltonian then becomes

$$G' = \begin{bmatrix} \frac{H_{11}+H_{22}}{2} & \frac{-Y+iX/2}{(3)^{\frac{1}{2}}} & \frac{Y+iX}{(6)^{\frac{1}{2}}} \\ \frac{Y-iX/2}{(3)^{\frac{1}{2}}} & \frac{H_{11}+H_{22}+4H_{33}}{6} & \frac{(Z/3)+iY\cos(\chi-2\eta)}{(2)^{\frac{1}{2}}} \\ \frac{Y-iX}{(6)^{\frac{1}{2}}} & \frac{(Z/3)-iY\cos(\chi-2\eta)}{(2)^{\frac{1}{2}}} & \frac{H_{11}+H_{22}+H_{33}}{3}-\Delta \end{bmatrix} \tag{28}$$

$$\Gamma' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{Y\sin(\chi-2\eta)}{(2)^{\frac{1}{2}}} \\ 0 & \frac{-Y\sin(\chi-2\eta)}{(2)^{\frac{1}{2}}} & 0 \end{bmatrix} \tag{29}$$

For k-vectors in the $x-y$ plane the quantity Y vanishes; for k in the (111) direction $\chi-2\eta=\pi$ so that for these cases the 6×6 matrix has been factored into two 3×3 matrices. In the general case further transformations are necessary. A brute-force method of making $\Gamma=0$ in general has been found but will not be given here since G becomes very complicated. It is not difficult to obtain the general secular equation as a product of two cubics directly from the matrix as given in eqs. (28) and (29). The general secular equation can be written

$$\begin{aligned} &H'_{11}H'_{22}H'_{33}+2H_{12}H_{23}H_{13}-H'_{11}H_{23}{}^2- \\ &-H'_{22}H_{13}{}^2-H'_{33}H_{12}{}^2-(\Delta/3)\{H'_{11}H'_{22}+ \\ &+H'_{11}H'_{33}+H'_{22}H'_{33}-H_{12}{}^2-H_{13}{}^2-H_{23}{}^2\}=0 \end{aligned} \tag{30}$$

$$H'_{ii}=H_{ii}+(\hbar^2/2m)k^2-E_k. \tag{31}$$

The symbols H_{ij} refer to the elements of the matrix given in eq. (6).

Along the simple directions (100) and (111) the secular equation factors into a quadratic term times a linear term. The linear term corresponds to a band which is unaffected by spin-orbit splitting and hence is parabolic within the limits of the approximations made in this derivation. In germanium the parabolic band is the upper valence band or heavy mass band.

All bands are parabolic in the limit of spin-orbit splitting very large or very small compared to the second-order $k\cdot p$ energy. The limit of $k\cdot p$ energy large compared to spin-orbit energy is not of real interest in germanium because higher-order perturbation terms which have been neglected here become important when the perturbation energy becomes of the order of the band gap. In silicon where $\Delta\simeq0{\cdot}04$ eV, the small Δ approximation may be useful for some purposes.

Calculations of the valence-band structure of germanium have been carried out using the values $L=-32{\cdot}0$ $(\hbar^2/2m)$, $M=-5{\cdot}30$ $(\hbar^2/2m)$, $N=-32{\cdot}4$ $(\hbar^2/2m)$ determined from cyclotron resonance constants as given by DKK. The value $\Delta=0{\cdot}28$ eV was used. After the calculations were completed, the comparison of theory and experiment indicated that $\Delta=0{\cdot}29$ eV was a better value for the spin-orbit splitting. The calculations have not been corrected for this change. The energy and slope have been calculated for all 3 bands and for the (100), (111), (110), and (Y) directions. The direction (Y) makes equal angles with the (100), (111), and (110) directions. The results are plotted in Fig. 1 as E vs. k^2. The slopes, $(2m/\hbar^2)dE/d(k^2)$, are plotted in Fig. 2. In a plot of E vs. k^2 the customary parabolas become straight lines. The surface of the Brillouin zone occurs at $k=0{\cdot}51$ atomic units in the (111) direction so that these calculations cover only a small part of the Brillouin zone.

For the (100) and (111) directions the heavy mass-band graphs are straight lines. For the (110) direction the heavy mass-band graph is nearly straight, the departure from a straight line is best seen in the graph of the slope $dE/d(k^2)$. For a larger range of k^2 the curvature of E vs. k^2 in the (110) direction is more noticeable.

For the (100) and (111) directions the light mass-band (band 2) bends up to become asymptotically parallel to band 1. In the (110) direction band 2 bends up appreciably but always remains steeper than band 1. The bending of band 2 is not of great significance to transport problems because the bend occurs at energies appreciably greater than kT. The effect is very noticeable in free carrier absorption, however, because the absorbing holes make vertical transitions from band 1 where the value of k^2 corresponding to thermal energies is much larger.

VOL. 1 1956/5

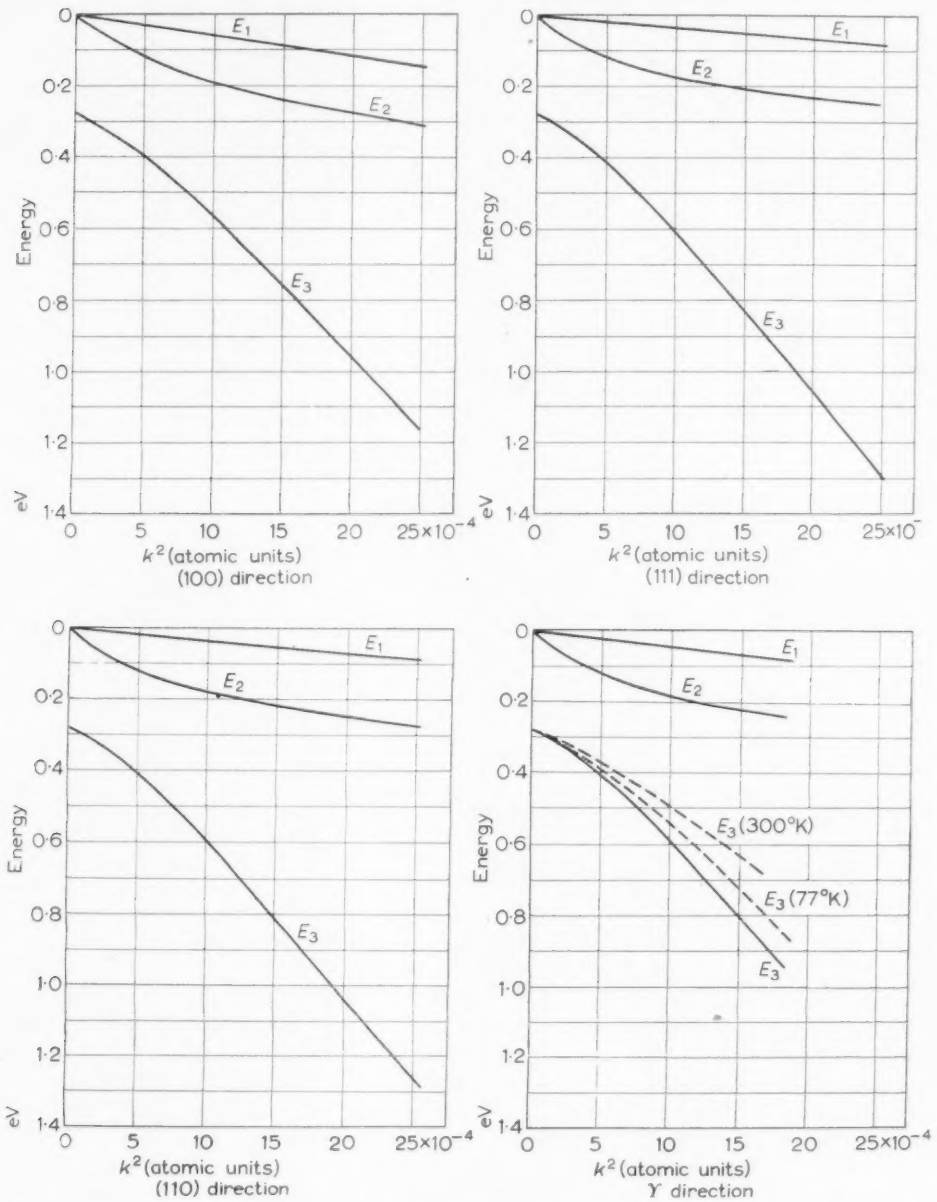


FIG. 1. E vs. k^2 for p-germanium. (Υ) direction shows correction to band 3 to give agreement with free carrier absorption at 300°K and 77°K.

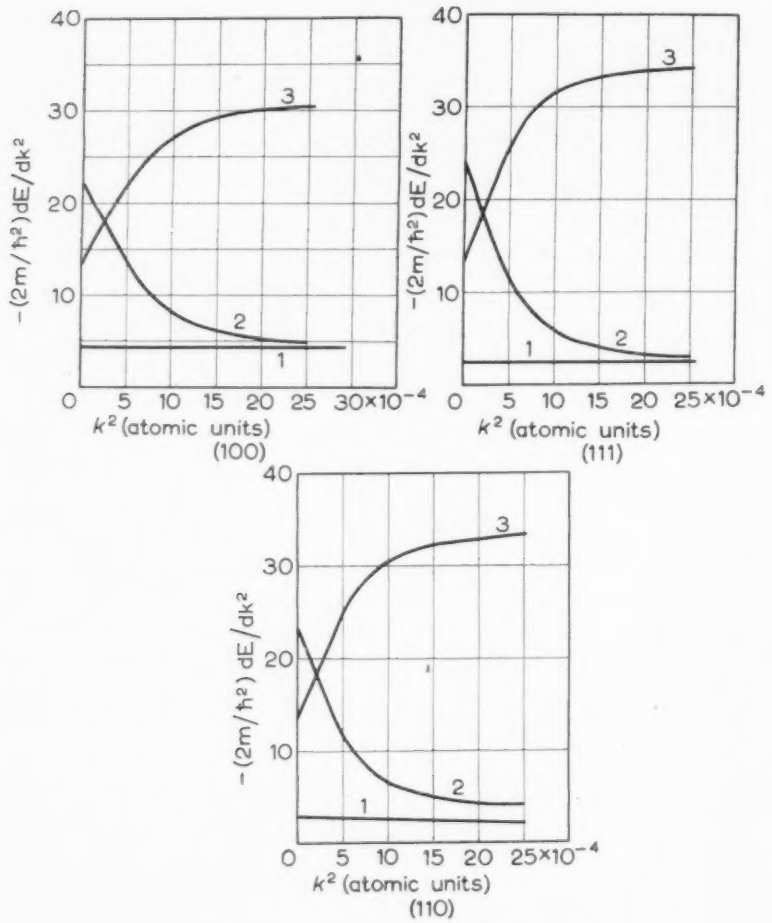


FIG. 2. Slope of E vs. k^2 for p-germanium.

The slope of band 3 increases with k^2. Band 3 is accurately isotropic for small k^2 and remains reasonably isotropic for large k^2.

The status of the energy bands in silicon is more uncertain owing to the ambiguity in sign of the cyclotron resonance constants. Furthermore, the limits of error of the constants are larger than in germanium, so that even for a given choice of signs there is considerable uncertainty in the effective mass of the heavy mass band. To illustrate this point, the value of the reciprocal effective mass in the limit of spin-orbit splitting small compared to the $k \cdot p$ energy for the (110) direction will be cited. The set of values $A = -4{\cdot}0$ $(\hbar^2/2m)$, $B = +1{\cdot}1$ $(\hbar^2/2m)$, $C = -4{\cdot}0$ $(\hbar^2/2m)$ gives a value $+0{\cdot}39$ for $(2m/\hbar^2)dE/d(k^2)$ in the (110) direction for large k. This is the choice of sign preferred by DKK. The other choice of sign $A = -4{\cdot}0$ $(\hbar^2/2m)$, $B = -1{\cdot}1$ $(\hbar^2/2m)$, $C = -4{\cdot}0$ $(\hbar^2/2m)$ gives a value of $-0{\cdot}71$ for $(2m/\hbar^2)dE/d(k^2)$. However, if we take the extreme values of A, B, C in a direction to make the reciprocal effective mass more negative with the DKK choice of sign the slope becomes $-0{\cdot}73$. A positive slope of $+0{\cdot}39$ could rule out the DKK choice of sign because the maxima in the valence band would occur away from $k = 0$. If the cyclotron resonance constants could be determined more

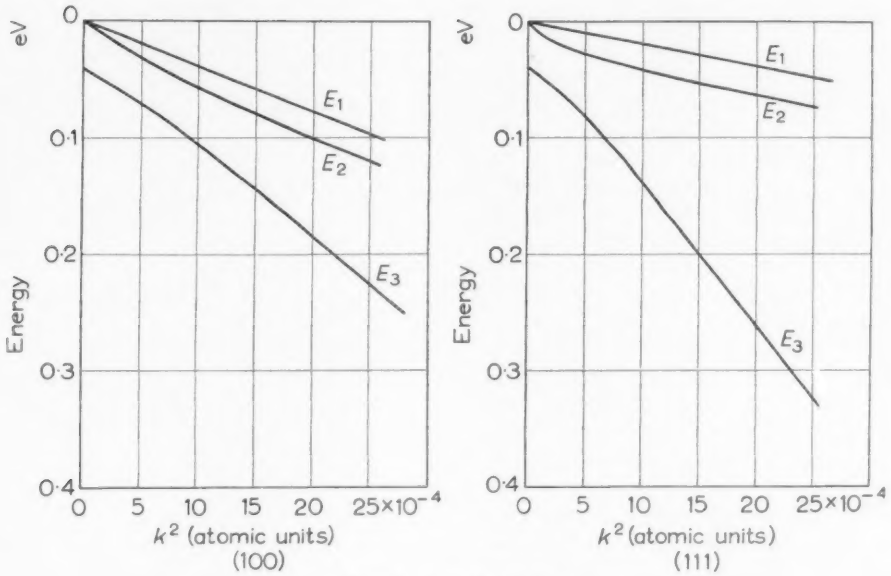


FIG. 3. E vs. k^2 for p-silicon.

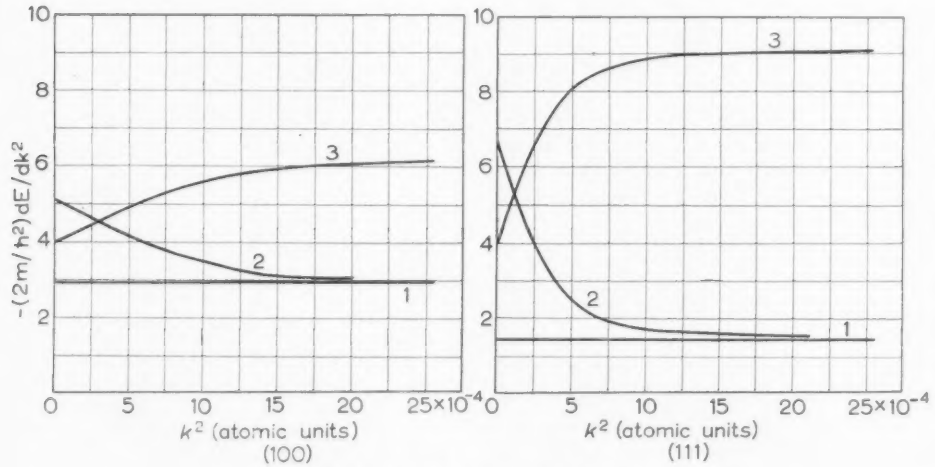


FIG. 4. Slope of E vs. k^2 for p-silicon.

accurately the correct choice of signs might be decidable.

The energy and slope of the valence bands in silicon have been calculated for the (100) and (111) directions. The results are given in Figs. 3 and 4. The values of the parameters used were $L = -7{\cdot}2$ $(\hbar^2/2m)$, $M = -3{\cdot}9$ $(\hbar^2/2m)$, $N = -7{\cdot}7$ $(\hbar^2/2m)$ corresponding to $A = -4{\cdot}0$ $(\hbar^2/2m)$, $B = -1{\cdot}1$ $(\hbar^2/2m)$, $C = -4{\cdot}0$ $(\hbar^2/2m)$. The choice of sign used here makes the silicon band structure qualitatively similar to germanium. This choice of sign is opposite to that given in DKK.

3. HIGHER-ORDER CORRECTIONS

These calculations of band structure have been based on the diagonalization of an interaction Hamiltonian between the triply degenerate valence wave functions at $k = 0$. The Hamiltonian was the sum of two parts: a first-order Hamiltonian of the spin-orbit interaction and a second-order Hamiltonian of the $k \cdot p$ perturbation. The technique of simultaneous diagonalization is clearly the best when both perturbation Hamiltonians are of the same order of magnitude. This procedure is formally indicated by making a perturbation expansion in powers of λ with the $(k \cdot p)$ perturbation multiplied by λ and the spin-orbit perturbation multiplied by λ^2. In nondegenerate perturbation theory the assignment of different powers of λ to different perturbations merely gives a rearrangement of the terms of the perturbation expansion. In degenerate perturbation theory, however, different relative powers in λ result in essentially different series. Convergence is always hastened by arranging powers of λ such that all perturbations act simultaneously to remove the degeneracy. If the degenerate Hamiltonians are of different orders of magnitude the procedure is still formally correct but unnecessary. Formulae for degenerate perturbation theory with two perturbations are developed in the appendix.

We make the convention that the degenerate states we are interested in are labeled by small subscripts j. Other states are labeled by capital subscripts J. For the case of a perturbation P of order λ where all matrix elements P_{ij} are zero and in addition a perturbation Q of order λ^2, we have eq. (A3) for the third-order energy, $E_i^{[3]}$, and eq. (A4) for the fourth-order energy, $E_i^{[4]}$, in the appendix. For our purposes P is the $k \cdot p$ perturbation which only connects states of opposite parity, Q is the spin-orbit interaction which only connects states of like parity. On parity grounds alone all third-order terms are zero and fourth-order terms 6 through 11 in eq. (A4) are zero. The nonzero terms will be repeated here.

$$E_i^{[4]} = \sum_K \frac{Q_{iK}Q_{Ki}}{E_i^0 - E_K^0} - E_i^{[2]} \sum_K \frac{P_{iK}P_{Ki}}{(E_i^0 - E_K^0)^2} + \sum_{K,L,J} \frac{P_{iL}P_{LK}P_{KJ}P_{Ji}}{(E_i^0 - E_L^0)(E_i^0 - E_K^0)(E_i^0 - E_J^0)} +$$

$$+ 2Re \sum_{K,L} \frac{P_{iL}P_{LK}Q_{Ki}}{(E_i^0 - E_L^0)(E_i^0 - E_K^0)} + \sum_{K,L} \frac{P_{iK}Q_{KJ}P_{Ji}}{(E_i^0 - E_K^0)(E_i^0 - E_J^0)}. \quad (32)$$

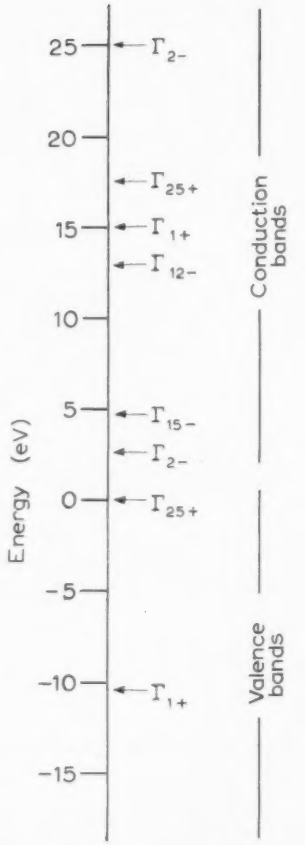


FIG. 5. Energy levels at $k = 0$ based on calculations by F. HERMAN. Reproduced from DRESSELHAUS, KIP, and KITTEL (reference 1).

The magnitude of these fourth-order terms will be discussed for germanium. The estimates refer to the (100) direction calculations but the magnitudes should be comparable for all directions. The energies of bands at $k = 0$ are given in Fig. 5, based on calculations by F. HERMAN[11]. The estimates of corrections to the present calculations from higher order terms are based on these $k = 0$

VOL. 1 1956/57

energies except that we assume that the Γ_2^- band is only 0·88 volts above the valence band as indicated by the optical absorption work of DASH and NEWMAN[12].

The second term is the most important term in eq. (32). The first factor, $E_i^{[2]}$, is the total second-order energy of the ith band, hence it contains both spin-orbit and $k \cdot p$ contributions. For small k, $E_i^{[2]}$ is approximately $+\Delta/3$ for bands 1 and 2 and $-2\Delta/3$ for band 3. The summation in term two is very similar to the second-order $k \cdot p$ energy of eq. (5) except that the resonance denominators are squared. The magnitude of the second term is then roughly $E_i^{[2]}(E_i^{[2]}(k) - E_i^{[2]}(0))/(E_i^0 - E_K^0)$ where E_K^0 refers to the band which contributes most to the summation. For band 1, $E_K^0 - E_i^0$ is about 5 volts since the first nonzero term comes from the Γ_{15} band. For bands 2 and 3, $E_K^0 - E_i^0$ is 0·88 volts at nitrogen temperature according to the optical absorption work of DASH and NEWMAN[12]. The lowest conduction band at $k = 0$ is assumed to be the Γ_2^- band. For small k, the second term causes an appreciable error in the calculation of effective masses. The effective mass calculations are about 2 per cent too large for band 1, 10 per cent too large for band 2, and 20 per cent too small for band 3. These estimates are based on the assumption that we have started with the correct values of L, M, N. Since the values used are experimentally derived they essentially include this correction so that we estimate that the correction to band 3 should be about 30 per cent and that bands 1 and 2 are reasonably accurate for small k.

The arguments above do not imply any error in the formula

$$E = Ak^2 \pm [B^2k^4 + C^2(k_x^2k_y^2 + k_y^2k_z^2 + k_z^2k_x^2)]^{\frac{1}{2}} \tag{33}$$

which has been given in DKK. This is the most general form for the energy surfaces and is perfectly rigorous for small k, except for the possibility that C^2 may be negative. The form follows from the requirements that the secular equation be a quadratic, that E be proportional to k^2, and that the energy surfaces have cubic symmetry. The error arises in relating the quantities L, M, N to A, B, C.

For large k the second term in eq. (32) becomes approximately $-(E_i^{[2]})^2/\Delta E$ where $\Delta E = 0{\cdot}88$ eV for band 3, 5 eV for band 1, and some average of 0·88 and 5 eV for band 2. For large k the second term in eq. (32) becomes proportional to k^4. The third term is also proportional to k^4 but appears to be less important than the second term because it involves states further from the valence band which have larger energy denominators.

The first term in eq. (32) is a small correction to the spin-orbit splitting at $k = 0$. The lowest state having a matrix element Q_{iK} with the valence band is the upper Γ_{25}^+ state, about 18 volts above the valence band. If Q_{iK} is of the order of the valence band splitting, the first term in eq. (32) amounts to about 2 per cent of the first order splitting, Δ.

Terms (4) and (5) in eq. (32) are proportional to k^2 times spin-orbit splitting hence they constitute an additional correction to the effective masses near $k = 0$. The largest contribution comes from term (5) and results from the spin-orbit splitting of the Γ_{15}^- band. If the Γ_{15}^- splitting is the same as the valence-band splitting, the fifth term will give a 6 per cent change in the effective mass of of the heavy-mass band. The per cent change will be smaller for bands two and three.

The error in the theoretical procedure should be about 10 times smaller for silicon than for germanium because of the smaller spin-orbit splitting. Unfortunately, the error in the cyclotron resonance constants is still appreciable for silicon and no experimental evidence is yet available which gives the value of the spin-orbit splitting accurately, though an estimate of 0·04 eV from the atomic splitting is probably not far off.

4. OPTICAL MATRIX ELEMENTS

The optical transition probability in free carrier absorption from band i to band j is proportional to $(e/mc)^2|\langle u_i|\boldsymbol{A} \cdot \boldsymbol{p}|u_j\rangle|^2$ in the case of direct transitions (i.e. transitions where k is conserved). The k of the light wave is negligible. The quantity $\boldsymbol{A}$ is the vector potential of the light. The momentum operator $\boldsymbol{p}$ has odd parity, hence it has no matrix elements between bands of like parity at $k = 0$. At $k = 0$, parity is a good quantum number in a cubic crystal. The $k \cdot p$ perturbation brings in a first order correction wave function of parity opposite to the $k = 0$ parity of the band. The operator $\boldsymbol{A} \cdot \boldsymbol{p}$ then has a matrix element between bands of like parity which is proportional to k.

The periodic part of the wave function, correct to first order in the $k \cdot p$ perturbation is

$$u = u_{0i} + \sum_J (\hbar/m) \frac{\langle u_{0J}|\boldsymbol{k}\cdot\boldsymbol{p}|u_{0i}\rangle}{E_i^0 - E_J^0} u_{0J}. \quad (34)$$

The matrix element of $A \cdot p$ between bands i and j is

$$(e/mc)\langle u_i|\boldsymbol{A}\cdot\boldsymbol{p}|u_j\rangle = (e\hbar/m^2c)\sum_J \frac{\langle u_{0i}|\boldsymbol{k}\cdot\boldsymbol{p}|u_{0J}\rangle\langle u_{0J}|\boldsymbol{A}\cdot\boldsymbol{p}|u_{0j}\rangle + \langle u_{0i}|\boldsymbol{A}\cdot\boldsymbol{p}|u_{0J}\rangle\langle u_{0J}|\boldsymbol{k}\cdot\boldsymbol{p}|u_{0j}\rangle}{E_i^0 - E_J^0}. \quad (35)$$

The right-hand side of eq. (35) is identical in form with the second-order $k \cdot p$ energy of eq. (5). Hence the matrix for the transition probability may be conveniently obtained by using the $k \cdot p$ part of the Hamiltonian matrix in eq. (21) (set $\Delta = 0$). Eq. (35) indicates that the substitution $k_l A_m + A_l k_m$ should be made for $k_l k_m$ and the matrix should be multiplied by the constant factor $(e/c\hbar)$. The resulting matrix F must then be further transformed to $F' = U^{-1}FU$ by the unitary matrix U which diagonalizes the Hamiltonian of eq. (21). The optical transition probability is proportional to the square of the absolute value of the appropriate matrix element of $\boldsymbol{A} \cdot \boldsymbol{p}$. In a cubic crystal this transition probability will be independent of the direction of $\boldsymbol{A}$ if we average over equivalent cubic directions. We define a quantity W_{ij}

$$W_{ij} = (\hbar^2/m^2)\overline{|\langle u_i|\boldsymbol{A}\cdot\boldsymbol{p}|u_j\rangle|^2}/A^2k^2 \quad (36)$$

where we have averaged over the direction of A and summed over the two final states which are Kramers degenerate. We have obtained expressions for the matrix elements W_{ij} in the (100), (111) and (110) directions which will be denoted by W_{ij}^{Δ}, W_{ij}^{Λ} and W_{ij}^{Σ} respectively.

For the (100) direction

$$W_{11}^{\Delta} = (4/3)M^2 \quad (37)$$

$$W_{12}^{\Delta} = (1/9)\{(2)^{\frac{1}{2}}S_{22}^{\Delta} + S_{23}^{\Delta}\}^2N^2 \quad (38)$$

$$W_{13}^{\Delta} = (1/9)\{S_{22}^{\Delta} - (2)^{\frac{1}{2}}S_{23}^{\Delta}\}^2N^2 \quad (39)$$

$$W_{22}^{\Delta} = (4/27)\{L+M+(S_{22}^{\Delta})^2L+(S_{23}^{\Delta})^2M + 2(2)^{\frac{1}{2}}(L-M)S_{22}^{\Delta}S_{23}^{\Delta}\}^2 \quad (40)$$

$$W_{23}^{\Delta} = (1/27)[9N^2+4(L-M)^2\{S_{22}^{\Delta}S_{23}^{\Delta} - (2)^{\frac{1}{2}}(S_{22}^{\Delta})^2+(2)^{\frac{1}{2}}(S_{23}^{\Delta})^2\}^2] \quad (41)$$

$$W_{33}^{\Delta} = (4/27)\{L+M+(S_{23}^{\Delta})^2L+(S_{22}^{\Delta})^2M - 2(2)^{\frac{1}{2}}(L-M)S_{22}^{\Delta}S_{23}^{\Delta}\}^2 \quad (42)$$

$$S_{22}^{\Delta} = \frac{h_{23}^{\Delta}}{\{(h_{23}^{\Delta})^2+(E_2^{\Delta}-h_{22}^{\Delta})^2\}^{\frac{1}{2}}} \quad (43)$$

$$S_{23}^{\Delta} = \frac{h_{22}^{\Delta}-E_2^{\Delta}}{\{(h_{23}^{\Delta})^2+(E_2^{\Delta}-h_{22}^{\Delta})^2\}^{\frac{1}{2}}} \quad (44)$$

$$h_{23}^{\Delta} = -\{(2)^{\frac{1}{2}}/3\}(L-M)k^2 \quad (45)$$

$$h_{22}^{\Delta} = (2L+M)k^2/3. \quad (46)$$

In the (100) direction the Hamiltonian of eq. (21) has $\Gamma = 0$. The matrix diagonalizing H is then

$$R^{\Delta} = 0; S^{\Delta} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & S_{22}^{\Delta} & S_{23}^{\Delta} \\ 0 & -S_{23}^{\Delta} & S_{22}^{\Delta} \end{bmatrix} \quad (47)$$

The above 3×3 matrices are components of the 6×6 matrix according to eq. (20). The elements S_{22}^{Δ}, S_{23}^{Δ} can also be expressed in terms of the energy E_3^{Δ}. The energy E_i^{Δ} is the energy of band i in the (100) direction. Similarly, E_i^{Λ} and E_i^{Σ} refer to energies in the (111) and (110) directions respectively.

For the (111) direction:

$$W_{11}^{\Lambda} = (4/27)(L+2M-N)^2 \quad (48)$$

$$W_{12}^{\Lambda} = (1/27)[2N^2+4(L-M)^2 + \{N^2-4N(L-M)\}\{(S_{23}^{\Lambda})^2+2(2)^{\frac{1}{2}}S_{22}^{\Lambda}S_{23}^{\Lambda}\}] \quad (49)$$

$$W_{13}^{\Lambda} = (1/27)[2N^2+4(L-M)^2 + \{N^2-4N(L-M)\}\{(S_{22}^{\Lambda})^2-2(2)^{\frac{1}{2}}S_{22}^{\Lambda}S_{23}^{\Lambda}\}] \quad (50)$$

VOL. 1 1956/57

$$W_{22}{}^{\Lambda}=(4/27)[L+2M+N\{(S_{22}{}^{\Lambda})^2-2(2)^{\frac{1}{2}}S_{22}{}^{\Lambda}S_{23}{}^{\Lambda}\}]^2 \quad (51)$$

$$W_{23}{}^{\Lambda}=(1/27)[\{2(L-M)+N\}^2+4N^2\{S_{22}{}^{\Lambda}S_{23}{}^{\Lambda}+(2)^{\frac{1}{2}}(S_{22}{}^{\Lambda})^2-(2)^{\frac{1}{2}}(S_{23}{}^{\Lambda})^2\}^2] \quad (52)$$

$$W_{33}{}^{\Lambda}=(4/27)[L+2M+N\{(S_{23}{}^{\Lambda})^2+2(2)^{\frac{1}{2}}S_{22}{}^{\Lambda}S_{23}{}^{\Lambda}\}]^2 \quad (53)$$

$$S_{22}{}^{\Lambda}=\frac{h_{23}{}^{\Lambda}}{\{(h_{23}{}^{\Lambda})^2+(E_2{}^{\Lambda}-h_{22}{}^{\Lambda})^2\}^{\frac{1}{2}}} \quad (54)$$

$$S_{23}{}^{\Lambda}=\frac{h_{22}{}^{\Lambda}-E_2{}^{\Lambda}}{\{(h_{23}{}^{\Lambda})^2+(E_2{}^{\Lambda}-h_{22}{}^{\Lambda})^2\}^{\frac{1}{2}}} \quad (55)$$

$$h_{23}{}^{\Lambda}=(2)^{\frac{1}{2}}Nk^2/3 \quad (56)$$

$$h_{22}{}^{\Lambda}=(L+2M+N)k^2/3. \quad (57)$$

In writing the matrix elements for the (110) direction we make the following auxiliary definitions.

$$S_{ij}{}^{\Sigma}=n_{ij}/(n_{1j}{}^2+n_{2j}{}^2+n_{3j}{}^2)^{\frac{1}{2}} \quad (58)$$

$$\begin{aligned} n_{1j}&=h_{12}{}^{\Sigma}h_{23}{}^{\Sigma}-h_{13}{}^{\Sigma}(h_{22}{}^{\Sigma}-E_j{}^{\Sigma})\\ n_{2j}&=h_{12}{}^{\Sigma}h_{13}{}^{\Sigma}-h_{23}{}^{\Sigma}(h_{11}{}^{\Sigma}-E_j{}^{\Sigma})\\ n_{3j}&=(h_{11}-E_j)(h_{22}{}^{\Sigma}-E_j{}^{\Sigma})-(h_{12}{}^{\Sigma})^2 \end{aligned} \quad (59)$$

$$\begin{aligned} h_{11}{}^{\Sigma}&=(L+M)k^2/2\\ h_{22}{}^{\Sigma}&=(L+5M)k^2/6\\ h_{12}{}^{\Sigma}&=Nk^2/2(3)^{\frac{1}{2}}\\ h_{13}{}^{\Sigma}&=Nk^2/(6)^{\frac{1}{2}}\\ h_{23}{}^{\Sigma}&=(L-M)k^2/3(2)^{\frac{1}{2}} \end{aligned} \quad (60)$$

where $E_j{}^{\Sigma}$ is the energy of band j in the (110) direction. The quantities $W_{ij}{}^{\Sigma}$ can then be written

$$W_{11}{}^{\Sigma}=(4/3k^4)\{\Delta(S_{31}{}^{\Sigma})^2+E_1{}^{\Sigma}\}^2 \quad (62)$$

$$W_{12}{}^{\Sigma}=(4/3k^4)\{\Delta S_{31}{}^{\Sigma}S_{32}{}^{\Sigma}\}^2+(1/9)(L-M)^2\{S_{11}{}^{\Sigma}S_{22}{}^{\Sigma}-S_{21}{}^{\Sigma}S_{12}{}^{\Sigma}+(2)^{\frac{1}{2}}(S_{11}{}^{\Sigma}S_{32}{}^{\Sigma}-S_{31}{}^{\Sigma}S_{12})\}^2+(1/18)N^2\{(2)^{\frac{1}{2}}(S_{11}{}^{\Sigma}S_{22}{}^{\Sigma}-S_{21}{}^{\Sigma}S_{12}{}^{\Sigma})+(S_{12}{}^{\Sigma}S_{31}{}^{\Sigma}-S_{11}{}^{\Sigma}S_{32}{}^{\Sigma}+(3)^{\frac{1}{2}}(S_{22}{}^{\Sigma}S_{31}{}^{\Sigma}-S_{21}{}^{\Sigma}S_{32}{}^{\Sigma})\}^2 \quad (63)$$

$$W_{13}{}^{\Sigma}=(4/3k^4)(\Delta S_{31}{}^{\Sigma}S_{33}{}^{\Sigma})^2+(1/9)(L-M)^2\{S_{11}{}^{\Sigma}S_{23}{}^{\Sigma}-S_{21}{}^{\Sigma}S_{13}{}^{\Sigma}+(2)^{\frac{1}{2}}(S_{11}{}^{\Sigma}S_{33}{}^{\Sigma}-S_{31}{}^{\Sigma}S_{13}{}^{\Sigma})\}^2+(1/18)N^2\{(2)^{\frac{1}{2}}(S_{11}{}^{\Sigma}S_{23}{}^{\Sigma}-S_{13}{}^{\Sigma}S_{21}{}^{\Sigma})+S_{13}{}^{\Sigma}S_{31}{}^{\Sigma}-S_{11}{}^{\Sigma}S_{33}{}^{\Sigma}+(3)^{\frac{1}{2}}(S_{23}{}^{\Sigma}S_{31}{}^{\Sigma}-S_{21}{}^{\Sigma}S_{33}{}^{\Sigma})\}^2 \quad (64)$$

$$W_{22}{}^{\Sigma}=(4/3k^4)\{\Delta(S_{32}{}^{\Sigma})^2+E_2{}^{\Sigma}\}^2 \quad (65)$$

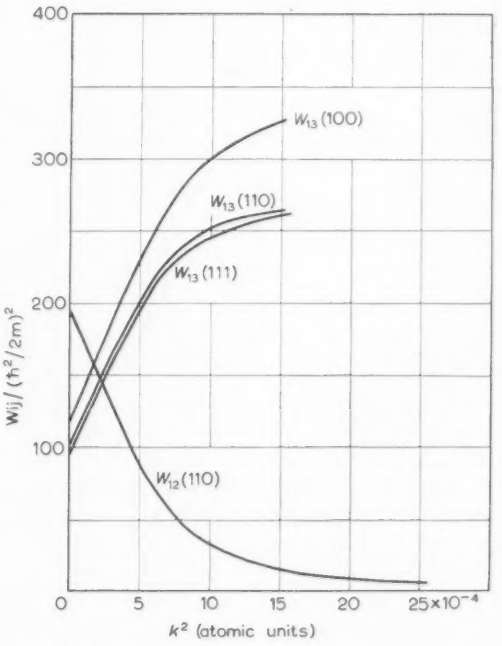


FIG. 6. Matrix elements for direct optical transitions in p-germanium.

$$W_{23} = (4/3k^4)\{\Delta S_{32}{}^{\Sigma} S_{33}{}^{\Sigma}\}^2 + \\ +(1/9)(L-M)^2\{S_{12}{}^{\Sigma} S_{23}{}^{\Sigma} - S_{22}{}^{\Sigma} S_{13}{}^{\Sigma} + \\ +(2)^{\frac{1}{2}}(S_{12}{}^{\Sigma} S_{33}{}^{\Sigma} - S_{32}{}^{\Sigma} S_{13}{}^{\Sigma})\}^2 + \\ +(1/18)N^2\{(2)^{\frac{1}{2}}(S_{12}{}^{\Sigma} S_{23}{}^{\Sigma} - S_{13}{}^{\Sigma} S_{22}{}^{\Sigma}) + \\ +S_{13}{}^{\Sigma} S_{32}{}^{\Sigma} - S_{12}{}^{\Sigma} S_{33}{}^{\Sigma} + \\ +(3)^{\frac{1}{2}}(S_{32}{}^{\Sigma} S_{23}{}^{\Sigma} - S_{33}{}^{\Sigma} S_{22}{}^{\Sigma})\}^2 \qquad (66)$$

$$W_{33}{}^{\Sigma} = (4/3k^4)\{\Delta(S_{33}{}^{\Sigma})^2 + E_3{}^{\Sigma}\}^2 \qquad (67)$$

Δ is the spin-orbit splitting.

The quantities $W_{13}{}^{\Delta}$, $W_{13}{}^{\Lambda}$, $W_{13}{}^{\Sigma}$, $W_{12}{}^{\Sigma}$ have been calculated. The results are plotted in terms of the dimensionless quantities $W_{ij}/(\hbar^2/2m)^2$ in Fig. 6. The value of $W_{23}{}^{\Delta}/(\hbar^2/2m)^2$ was calculated to be 560 for $k = 0$.

5. FREE-CARRIER ABSORPTION

The free-carrier absorption of p germanium has been discussed by KAHN[9]. In his treatment he made the parabolic band approximation. We shall extend his analysis using the band calculations and matrix elements computed in this paper.

The probability per unit time w for absorbing a photon is

$$w = (2\pi/\hbar)H^2\rho(E) \qquad (68)$$

where ρ is the density of states per unit energy. The transition probability can be related to the absorption constant, α, by the equation

$$\alpha = 2\pi\hbar^2 cw/nEA^2 \qquad (69)$$

where E is the energy of the absorbed quantum and n is the index of refraction. According to eq. (36) the matrix element squared is given by

$$H_{ij}{}^2 = (e/c\hbar)^2 A^2 k^2 W_{ij}. \qquad (70)$$

The density of states is determined by the number of possible final states in the energy range $d(E_i - E_j)$ where a hole is to be taken from band i to band j by a direct transition. The number of one electron states per unit volume in this range in band i is

$$k^2/\pi^2[(dE_i/dk) - (dE_j/dk)] \qquad (71)$$

for the case of spherical energy bands. To each occupied state in band i there corresponds a final state of the system in which the hole in band i has been transferred to a state of the same k vector in band j. The density of final states is just the density of states in (71) multiplied by the occupation probability. For Boltzmann statistics the density of states is

$$\rho(E) = \frac{k^2 p e^{-E_i/kT}}{\pi^2 N_v |(dE_i/dk) - (dE_j/dk)|} \qquad (72)$$

$$N_v = 2(kT/2\pi\hbar^2)^{3/2}(\overline{m}_1{}^{3/2} + \overline{m}_2{}^{3/2}) \qquad (73)$$

where E_i and E_j are the energies in the bands i and j between which transitions are made. The concentration of holes is p; $\overline{m}_1$, $\overline{m}_2$ are the average masses of the two upper valence bands. A density of states factor of two for spin is included for band i but not for band j in eq. (72) since the latter factor of two has already been accounted for in computing W_{ij} by summing over the two Kramers degenerate bands of the final state.

The equation for the absorption constant is then

$$\alpha_{ij} = \frac{2e^2 k^3 p W_{ij}(e^{-E_i/kT} - e^{-E_j/kT})}{c\hbar n N_v (E_i - E_j)[(dE_i/dk^2) - (dE_j/dk^2)]}. \qquad (74)$$

Induced emission and absorption are both included. The above formula applies only to the case of spherical energy bands and Boltzmann statistics.

Band 1 to Band 3 Transitions

The free-carrier absorption has been calculated for the (100), (111), (110), and (Ψ) directions separately as if one had spherical bands. A crude numerical integration is then effected by weighting the directions appropriately. The weighting was made by considering the surface of a sphere in k space and assigning a small element of area to whichever of the four computed directions it was closest. The weightings were then made proportional to the area and were normalized. The following approximate weights were determined: (100), 0·09: (111), 0·16: (110), 0·22: (Y), 0·53.

The computations for the (100), (111), and (110) directions include explicit evaluation of all quantities given in eq. (74). For the (Y) direction, however, only the Boltzmann factor and the energy appropriate to (Y) were used. The matrix elements and values of $dE/d(k^2)$ were taken to be the same as

VOL. 1 1956/57

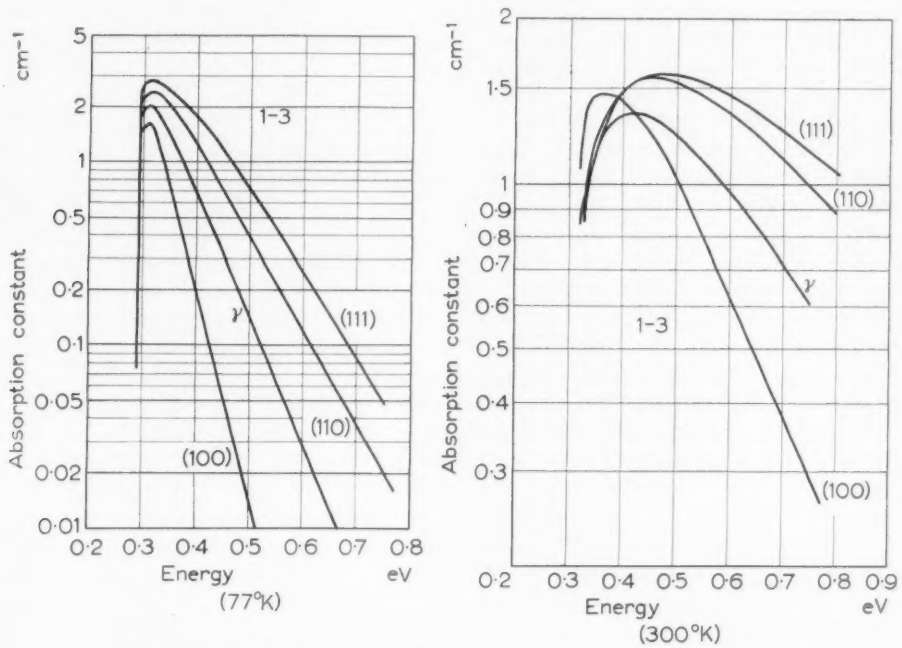


FIG. 7. Theoretical free-carrier absorption in *p*-germanium with 6×10^{15} holes.

for the (110) direction. The computations of absorption constant for the four directions are given in Fig. 7 for 77°K and 300°K. The weighted averages are compared with the experimental data of R. NEWMAN[10] in Fig. 8.

For both 77°K and 300°K the peak theoretical values of absorption constant are about a factor of 1·2 high. Since the theory contains no adjustable parameters this agreement is quite good. However the high-energy slopes show a significant discrepancy between theory and experiment for both 300°K and 77°K. The source of the disagreement is thought to be the neglected higher-order terms in the perturbation expansion which will have the effect of bringing band 3 closer to bands 1 and 2. The effect of higher-order terms on band 3 will be much larger than the effect on bands 1 and 2. The experimental curve can be used to obtain a rough estimate of the correction which should be made to band 3. The theoretical curve is first divided by a factor of 1·2 so that both curves have the same peak absorption. Then the horizontal energy displacement which is needed to make the two curves superpose is taken to be the vertical energy displacement which should be added to band 3. A given absorption constant defines two values of E_1-E_3, a value A for the experimental curve and a value B for the theoretical curve. The correction $A-B$ to band 3 is to be made at the point in k space where the band separation E_1-E_3 is equal to B. This procedure assumes that the correction to band 3 is isotropic; it also assumes that the correction to band 3 will not alter the absorption computed for a given point in k space. The latter assumption is reasonable at 77°K but not too good at 300°K.

The correction to band 3 is made for both 77°K and 300°K and is plotted in Fig. 1 for the (Y) direction. Ideally, of course, the two corrections should coincide but actually there is a large discrepancy. A number of reasons which may contribute to the discrepancy can be cited.

The theoretical curve at nitrogen temperature is very sensitive to the cyclotron resonance constants because of the Boltzmann factor in band 1. The magnitude of the uncertainty in the absorption constant appears to be about a factor of 2 at 0·6 eV according to the limits of error quoted by

DKK. The slope at nitrogen temperatures may be influenced by phonon induced indirect transitions. The electron-phonon interaction will mix a small amount of large k vector wave function into the small k states. The probability of making a high-energy transition due to this admixture of k vector is proportional to H^2 (electron-phonon)$/E_1^2$ where E_1 is the large k energy in the upper valence band. For large E_1 this indirect transition probability will eventually exceed the direct transition probability which is multiplied by a Boltzmann factor $e^{-E_1/kT}$. A rough estimate of this effect indicates that it may not be negligible. However, it would appear that such an effect should cause a change in slope of absorption vs. energy for large energies which is not observed experimentally. This effect would be in a direction to increase the indicated correction to band 3 at nitrogen temperature.

The theoretical absorption at room temperature varies by less than a factor of 2 from 0·4 eV to 0·7 eV. This fact makes the computed correction to band 3 very sensitive to the value of the optical matrix element as a function of k. The computed matrix elements are influenced by the neglected higher-order terms just as are the band shapes. The direction of this effect is to reduce the indicated correction to band 3 at room temperature.

The room-temperature correction to band 3 would be changed somewhat by taking into account self-consistently the effect on the absorption of the correction itself. The direction of this effect is to increase the indicated correction.

A temperature dependence of the band E_3 is indicated by the fact that the band gap at $k = 0$ is significantly temperature dependent between 77°K and 300°K as shown by the optical absorption studies of DASH and NEWMAN[12]. This effect would decrease the effective mass of band 3 by about 10 per cent in going from nitrogen to room temperature and is in the opposite direction to the apparent change inferred from the present method of correcting the E_3 band.

In summary, no accurate method of determining the band E_3 has been found. The required alterations to theory are of the order of magnitude ($\sim$ 30 per cent) to be expected from estimates of higher-order terms. Although 300°K and 77°K corrections do not agree, they appear to be reconcilable in view of the uncertainties mentioned above. The correction to E_3 deduced from 77°K data should be most accurate, particularly if the accuracy of the cyclotron resonance constants increases.

Band-1 to Band-2 Transitions

The theoretical calculation of the free-carrier absorption at 300°K for transitions from band 1 to band 2 is given in Fig. 8 for the (110) direction. The decrease in slope of E vs. k^2 in band 2 for large k^2 brings the absorption down sharply before 0·23 eV where the experimental curve shows a minimum. The effect is even more marked for the (100) and (111) directions. In these directions, bands 1 and 2 become parallel asymptotically in the approximation of our calculations. The minimum separation of bands 1 and 2 is $(2/3)\Delta$ or 0·19 eV for both the (100) and (111) directions. Hence the absorption goes to zero at this point. The absolute magnitude of the theoretically computed absorption agrees quite well with the experimental value.

Band-2 to Band-3 Transitions

The approximation of direct transitions breaks down in the analysis of transitions between bands 2 and 3. Since the slopes are equal at $k^2 = 4\times10^{-4}$ atomic units the density of states in the spherical approximation becomes infinite at this point. The absorption approaching the point at which the slopes are equal increases as $(E-E_0)^{-\frac{1}{2}}$. The integrated absorption $\int \alpha\, dE$ remains finite. Any scattering mechanism such as phonons or impurities which makes the k vector of an electron uncertain will broaden the infinite narrow spike associated with direct transitions and produce a finite peak.

A source of broadening apart from indirect transitions is the fact that the surface in k space for which the slopes in bands 2 and 3 are equal does not coincide with the surface of constant energy separation. The maximum broadening to be expected from this source is only $1{\cdot}5\times10^{-4}$ eV which is far less than the experimental width of about 0·02 eV.

In the range of 10^{15} carriers/cc the free-carrier absorption is independent of impurity concentration[10] so that the broadening of the α_{23} band is not caused by impurity scattering. The most likely source of broadening appears to be lattice scattering.

VOL. 1 1956/57

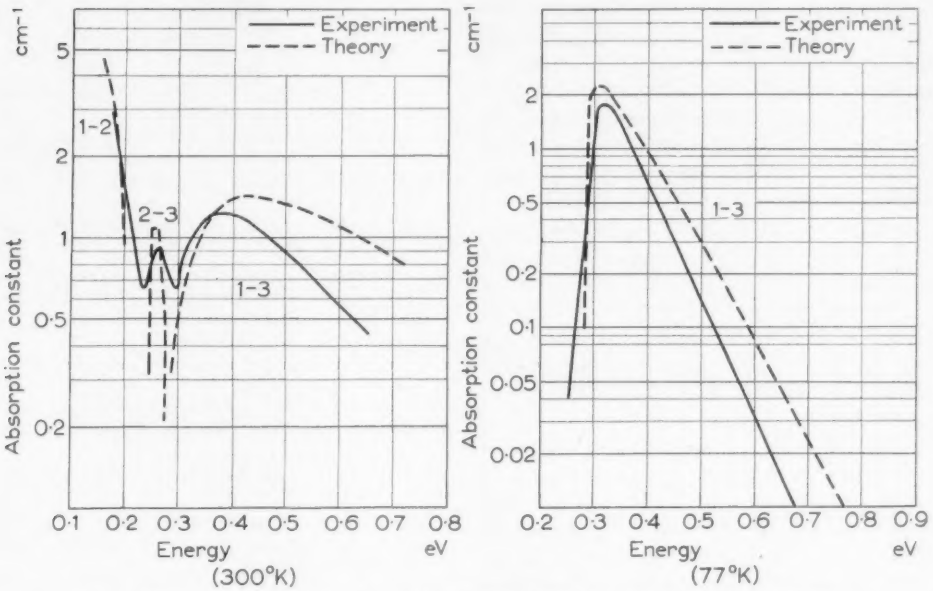


FIG. 8. Theoretical free-carrier absorption compared with experiment for 6×10^{15} holes.

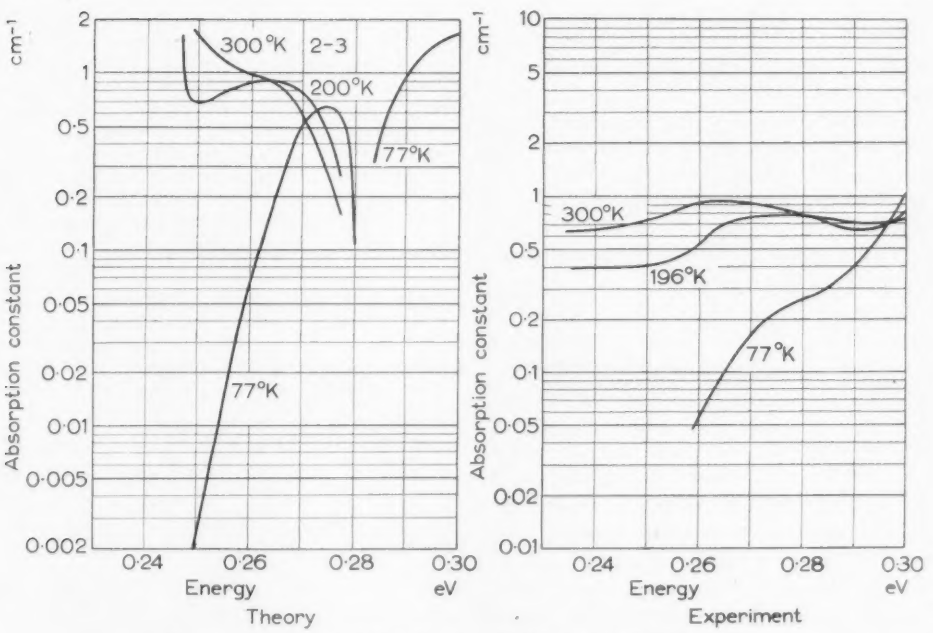


FIG. 9. Free-carrier absorption for band-2 to band-3 transitions.

The theoretical calculations of free-carrier absorption for band-2 to band-3 transitions are shown in Fig. 9. The calculations were made for the (100) direction; the matrix element W_{23} was calculated for $k = 0$ only, since the variation with k is not very important. The assumption of direct transitions causes a rapid rise in absorption on the low-energy side of the curve. This rise is shown in the 200°K and 300°K curves. At 77°K the absorption falls to very low values before it rises again because of the infinite density of states. At 77°K the shape of the curve is not very sensitive to how the infinite, narrow peak is broadened. At 300°K the shape of the curve is entirely dependent on how the broadening is treated so that the direct transition approximation is of no value here.

The experimental α_{23} band is also shown at three temperatures in Fig. 9. The theoretical peak absorption at 200°K agrees roughly with the experimental values. At 77°K the α_{23} band has moved so close to the α_{13} band that the bands are not clearly resolved experimentally, presumably because of phonon broadening. It is difficult to compare theory and experiment for this case, although the absorption appears to be significantly lower than expected from theory.

For completeness the α_{23}-band calculations are included in Fig. 8. The absorption is arbitrarily clipped off at a peak value a factor of 1·2 above the experimental value.

A comparison of theory and experiment indicates that the spin-orbit splitting should have been taken as 0·29 eV. The α_{23} peak at 200°K, the α_{13} peak at 77°K, and the trough between the α_{13} and α_{23} peaks at 200°K all show this clearly. The most significant indication of the spin-orbit splitting is the trough between the α_{13} and α_{23} peaks at 200°K since the trough should occur quite accurately at the value Δ. The calculations have not been corrected for this revised estimate of the splitting.

Acknowledgements—The author is indebted to A. W. OVERHAUSER for many valuable discussions and to J. BARDEEN, H. BROOKS and H. EHRENREICH for useful suggestions. He is also grateful to R. NEWMAN for the use of his experimental data and for helpful comments. He wishes to express his appreciation to Miss ANN WARNER for performing the numerical calculations.

REFERENCES

1. DRESSELHAUS G., KIP A. F. and KITTEL C. *Phys. Rev.* **98,** 368 (1955).
2. LAX B., ZEIGER H. J., DEXTER R. N. and ROSENBLUM E. S. *Phys. Rev.* **93,** 1418 (1954).
3. DEXTER R. N., ZEIGER H. J. and LAX B. *Phys. Rev.* **95,** 557 (1954).
4. BRIGGS H. B. and FLETCHER R. C. *Phys. Rev.* **91,** 1342 (1953).
5. KAISER W., COLLINS R. J. and FAN H. Y. *Phys. Rev.* **91,** 1380 (1953).
6. SHOCKLEY W. *Phys. Rev.* **78,** 173 (1950).
7. ADAMS E. N. Report CML-TN-P8 (September 1954), Chicago Midway Laboratories (unpublished).
8. OVERHAUSER A. W. (private communication).
9. KAHN A. H. *Phys. Rev.* **97,** 1647 (1955).
10. NEWMAN R. (to be published in *Phys. Rev.*).
11. Reproduced from DRESSELHAUS G., KIP A. F. and KITTEL C. (reference one).
12. DASH W. C. and NEWMAN R. *Phys. Rev.* **99,** 1151 (1955).
13. CONDON E. U. and SHORTLEY G. H. *The Theory of Atomic Spectra* p. 34, Cambridge University Press (1953).

VOL. 1 1956/57

APPENDIX: DEGENERATE PERTURBATION THEORY

Two Perturbations of Different Order

If two perturbations are of different orders of magnitude they should clearly be assigned different powers of the expansion parameter λ in making a perturbation expansion. In the nondegenerate case, a different assignment of powers of λ merely results in a reordering of the terms of the perturbation series. In the case of degenerate perturbation theory, different powers of λ result in essentially different terms in the perturbation expansion.

Let us consider the particular case of two perturbations, P and Q, where P removes degeneracy in second order and Q removes degeneracy in first order. If $\sum_K P_{iK}P_{Kj}(E_i^0 - E_K^0)^{-1}$ is of the order of Q_{ij}, then clearly we should write $\lambda P + \lambda^2 Q$ in making a perturbation expansion. Then both perturbation matrices will be diagonalized together in determining the correct unperturbed function set. Convergence is never impaired by this procedure even when $\sum_K P_{iK}P_{Kj}(E_i^0 - E_K^0)^{-1}$ is actually much smaller than Q_{ij}.

We give the formulae for the energies up to fourth order for this case assuming $P_{ii} = 0$ so that the first-order energy vanishes.

$$E_i^{[1]} = 0 \tag{A1}$$

$$E_i^{[2]} = Q_{ii} + \sum_K \frac{P_{iK}P_{Ki}}{E_i^0 - E_K^0} \tag{A2}$$

$$E_i^{[3]} = 2Re \sum_K \frac{Q_{iK}P_{Ki}}{E_i^0 - E_K^0} + \sum_{K,J} \frac{P_{iJ}P_{JK}P_{Ki}}{(E_j^0 - E_J^0)(E_i^0 - E_K^0)} + \sum_{J,K,L} \frac{P_{iL}P_{LK}P_{KJ}P_{Ji}}{(E_i^0 - E_L^0)(E_i^0 - E_K^0)(E_i^0 - E_J^0)} + 2Re \sum_{K,L} \frac{P_{iL}P_{LK}Q_{Ki}}{(E_i^0 - E_L^0)(E_i^0 - E_K^0)} + \sum_{J,K} \frac{P_{iK}Q_{KJ}P_{Ji}}{(E_i^0 - E_K^0)(E_i^0 - E_J^0)} + \tag{A3}$$

$$E_i^{[4]} = \sum_K \frac{Q_{iK}Q_{Ki}}{E_i^0 - E_K^0} - E_i^{[2]} \sum_K \frac{P_{iK}P_{Ki}}{(E_i^0 - E_K^0)^2} + \sum_{k,J,K,L,M}{}' \frac{P_{iL}P_{LM}P_{Mk}P_{kK}P_{KJ}P_{Ji}}{(E_i^0 - E_L^0)(E_i^0 - E_M^0)(E_i^0 - E_K^0)(E_i^0 - E_J^0)\,(E_i^{[2]} - E_k^{[2]})} + 2Re \sum_{k,J,K,L}{}' \frac{Q_{iL}P_{LK}P_{kK}P_{KJ}P_{Ji}}{(E_i^0 - E_L^0)(E_i^0 - E_K^0)(E_i^0 - E_J^0)(E_i^{[2]} - E_k^{[2]})} + 2Re \sum_{k,J,K,L}{}' \frac{P_{iL}Q_{Lk}P_{kK}P_{KJ}P_{Ji}}{(E_i^0 - E_L^0)(E_i^0 - E_K^0)(E_i^0 - E_J^0)(E_i^{[2]} - E_k^{[2]})} + 2Re \sum_{k,K,L}{}' \frac{P_{iL}Q_{Lk}P_{kK}Q_{Ki}}{(E_i^0 - E_L^0)(E_i^0 - E_K^0)(E_i^{[2]} - E_k^{[2]})} + \sum_{k,K,L}{}' \frac{Q_{iL}P_{Lk}P_{kK}Q_{Ki}}{(E_i^0 - E_L^0)(E_i^0 - E_K^0)(E_i^{[2]} - E_k^{[2]})} + \sum_{k,K,L}{}' \frac{P_{iL}Q_{Lk}Q_{kK}P_{Ki}}{(E_i^0 - E_L^0)(E_i^0 - E_K^0)(E_i^{[2]} - E_k^{[2]})}. \tag{A4}$$

Capital indices J, K, L index states not degenerate with the state i under consideration. Indices j, k, l refer to states which are degenerate with i in the unperturbed problem. Primes on the summation signs indicate that the term $i = j$ is omitted. The unperturbed function set has been determined by the requirement

$$\sum_K \frac{P_{iK}P_{Kj}}{E_i^0 - E_K^0} + Q_{ij} = 0 \quad i \neq j.$$

Similar formulae are discussed in CONDON and SHORTLEY[13].

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 100–116.

THERMAL ACCEPTORS IN GERMANIUM*

HARRY LETAW, Jr.

Electrical Engineering Research Laboratory, University of Illinois, Urbana, Illinois†

(*Received* 27 *February* 1956)

Abstract—Thermal acceptors in germanium are identified as vacancies with an energy of formation 2 eV. The annealing of thermally introduced acceptors in germanium is described by a mechanism involving the formation of divacancies as the first step. The activation energy for the formation of a divacancy is 1·7 eV, consisting of a contribution of 0·96 eV from the activation energy for the diffusion of a vacancy and 0·7 eV from the long-range repulsion of two vacancies. The energy of formation and the binding energy of a divacancy are 1·1 and 2·9 eV, respectively. The diffusion coefficient of a divacancy possibly may be represented by $D_{VV} = 0{\cdot}1 \exp(-2{\cdot}2/kT)\ \mathrm{cm^2 sec^{-1}}$. At dry-ice temperatures, a vacancy accepts one electron and a divacancy, four, while at room temperature these are one and two, respectively. The second step of the annealing process involves the formation of clusters of four, the activation energy for the formation of which is 2·8 eV. Further annealing may take place through the formation of clusters of higher order. The behavior of the model is compared with experiments reported by others, and it is shown to account for the slow rate of annealing as well as the initial increase in acceptor density often observed during annealing.

VOL.
1
1956/5

1. INTRODUCTION

SINGLE-CRYSTAL germanium is a material which may be obtained relatively free of imperfections. A large number of investigators have studied its behavior, both in the pure state and following the addition of various classes of defects. We shall be interested primarily in the phenomenon of thermal conversion and related effects and, in even more detail, the annealing processes which permit non-equilibrium distributions of thermal defects to be annihilated. In order to form a clear idea of the experimental results obtained to this time, we shall begin with a critical review of the literature. It will first be necessary to separate the effects attributable to thermally introduced lattice defects, i.e., vacancies, interstitials, or dislocations, from those caused by impurity atoms (notably copper) either accidentally or intentionally added to the lattice.

It will be important to bear in mind during this review the fact that the rapid diffusion of vacancies and the slow progress of annealing (both discussed below) are not consistent with a simple mechanism for annealing. We shall account for the experimental observations by proposing that the very mobile vacancies may encounter one another and form highly stable, relatively slowly moving divacancies. Annealing, then, will be considered in the light of the properties of these and other vacancy clusters of higher order.

For some years, it has been well known that *n*-type germanium rapidly quenched from high temperatures tends to be converted to *p* type.[1] This change in conductivity type has been found to be reversible by annealing at somewhat lower temperatures. It was subsequently discovered that the thermally introduced acceptor state was 0·03 eV above the valence band.[2] FULLER, THEUERER, and VAN ROOSBROECK[3] measured the movement of a *p-n* junction into the surface of an *n*-type specimen of germanium heated in a helium atmosphere and concluded that the diffusion coefficient of acceptors moving in from the surface could be represented by $D = 0{\cdot}02 \exp(-0{\cdot}52/kT)\ \mathrm{cm^2\ sec^{-1}}$ (all energies are given in electron volts). These authors also noted that the equilibrium acceptor density varied with temperature in such a way that an energy of 1·2 to 1·4 eV may be

* Supported in part by the U.S. Air Force through the Office of Scientific Research of the Air Research and Development Command and in part by the Office of Naval Research.

† Present address: Research Division, Raytheon Manufacturing Company, Waltham 54, Massachusetts.

associated with the formation (or solution) process. TAYLOR[4] measured the energy of formation of thermal defects and found their density to be proportional to $\exp(-G/kT)$, where T is the temperature before quenching and $G = (1{\cdot}85-5{\cdot}43 \times 10^{-4}T)$ eV. DUNLAP,[5] in similar measurements, determined the energy of formation to be in the neighborhood of 1 eV.

FULLER and STRUTHERS[6] postulated that the acceptors responsible for thermal conversion are copper atoms which have diffused into the specimen from the surface. Experiments with radioactive copper indicated that the acceptor density in a given region was proportional to copper density. The equilibrium density of acceptors as a function of temperature was expressed by $N = 6{\cdot}3 \times 10^{21} \exp(-1{\cdot}2/kT)$ cm^{-3}. Crystals converted by the addition of copper could be reconverted by annealing at 500°C. The proposed mechanism for the cycle invoked the precipitation of copper as a separate phase upon annealing with its redispersal following upon heating at higher temperatures. Rapid quenching permitted the electrical properties of the crystal containing dispersed copper to be measured. The foregoing investigation was essentially confirmed by the proof that under a variety of conditions, growing crystals do not convert from *n*-type to *p*-type.[7] Conversion was achieved, however, by placing copper on the surface of an uncut crystal and heating.

A systematic study of thermal conversion was carried out by GOLDBERG.[8] Germanium specimens were quenched from high temperatures by being dropped from a furnace into an oil bath in a time of about 0·3 sec. The energy of formation of thermally introduced acceptors was found to be 1·49 eV. It was assumed that the acceptors were vacancies. ESAKI[9] carried out experiments which permitted him to determine the equilibrium density of acceptors as a function of temperature, $N = 2{\cdot}5 \times 10^{21} \exp(-1{\cdot}2/kT)$ cm^{-3}. Under the assumption that the acceptors are produced at the surface and then diffuse into the interior of the specimen, the diffusion coefficient was estimated as $D = 300 \exp(-1{\cdot}6/kT)$ cm^{2}sec^{-1}.

Further evidence for the identification of thermal acceptors as copper was presented by MORIN and MAITA[10] from studies on presumably pure germanium and on copper-doped germanium, each quenched from high temperatures. Acceptor levels were found in germanium pulled from a melt containing copper, in germanium into which copper has been diffused by heat treatment, and in heat-treated germanium to which no copper had been intentionally added. In each of the three cases, the levels were found to lie about 0·040 eV above the valence band. Lower-lying acceptor levels were detected in some specimens.

A series of experiments was carried out for the purpose of measuring the thermal acceptor density in the absence of copper. The first of these utilized specimens of germanium which were acid etched and then rinsed in double-distilled water.[11] Following this, they were soaked in a strong aqueous potassium cyanide solution which forms soluble complexes with copper and other metal ions.* LOGAN performed an experiment similar to that of GOLDBERG[8] with a quenching time of 0·1 sec. Using the expression $N = N_S \exp(G/kT)$ and setting the pre-exponential factor equal to the density of lattice sites in germanium, he found that his data were best fitted by 1·8 eV for the *average free energy* of acceptor production. FINN[12] was able to decrease the number of acceptors introduced into quenched germanium by high-temperature vacuum annealing rather than by annealing in helium at 1 atmos. pressure. With the use of Cu^{64}, he estimated that 96 per cent of the copper previously diffused into the body of the specimen was evaporated during a vacuum heating. The loss of copper corresponded within experimental error to the change in acceptor concentration. Further investigations which supported the idea that thermal conversion is attributable to impurities diffusing from the surface were carried out by SEILER *et al.*[13] and MAESEN, PENNING, and WIERINGEN.[14] These experiments dealt, in the main, with large single crystals.

In an investigation of the solubility of germanium in a number of metals, a portion of the germanium crystal was covered by an alloy of germanium and a particular (unspecified) metal and heated to high temperatures for various lengths of time.[15] Although no donor impurity was believed to be present in the alloy, that portion of the crystal

* This procedure has been confirmed in this laboratory by W. M. PORTNOY who used specimens plated with radioactive Cu^{64}. It was found that a vigorous etching with *CP*-4 removed most of the activity within the order of 10 sec.

immediately beneath the alloy melt remained unconverted even though the remainder of the crystal was converted to *p*-type. In another experiment, an *n*-type crystal was immersed in a melt containing germanium and the same metal and was held at 900°C for several hours. The melt was rapidly quenched and the germanium specimen when removed was found to be *n*-type throughout. A similar crystal was thermally converted by heating at the same temperature for several minutes in the absence of the melt. It was then heated for 15 minutes in the melt and quenched. The crystal became *n*-type throughout. An *n*-type crystal was heated in both sodium chloride and boric oxide melts at 900°C and was converted to *p*-type in each case. Reconversion was accomplished in the alloy melt. If analytically determinable amounts of copper were added to the melt, the germanium became *p*-type. These experiments were interpreted by TROUSIL as indicating that so long as the crystal was in contact with molten germanium, Schottky defects diffused out into the melt and were there annihilated. He pointed out that an atom removed from the surface was more easily replaced by an atom from the melt than by one from the bulk. Thus, no Schottky defects were created in this environment. He stated that both copper and Schottky defects act as acceptors.

MAYBURG extended an earlier investigation[16] to a study of the mechanism of annealing of thermal acceptors quenched into germanium.[17] The density of defects introduced was representable as a function of quenching temperature by $N = 3\times10^{23}\exp(-2{\cdot}01/kT)\ \text{cm}^{-3}$. At annealing temperatures higher than 516°C, the nonequilibrium acceptor density decreased monotonically with time. A semilogarithmic plot of acceptor density against time was found to be nonlinear; thus, the process could not be described by a simple model involving only the loss of vacancies at the surface. At lower temperatures, the acceptor density increased with annealing time until a maximum value was attained. Upon further annealing, acceptors were gradually lost from the specimen. The physical mechanism called upon to describe this behavior utilized Frenkel defects in which (negatively charged) vacancies were less mobile than (positively charged) interstitials at temperatures below 516°C. It was supposed that at the lower temperatures, interstitials, migrating rapidly to dislocation lines, became trapped and discharged and were later annihilated by vacancies. Above 516°C, it was assumed that sufficient thermal energy was imparted to interstitials so that they could not be trapped; hence, they recombined directly with vacancies.

LOGAN and SCHWARTZ defined the conditions under which thermal conversion does not occur.[18] Specimens were heated in vacuum between molybdenum electrodes. The possibility of plastic deformation was minimized. Specimens were cooled from temperatures in the neighborhood of 875°C with no observable change in either resistivity or lifetime so long as the rate of cooling was no greater than $100°\text{C}-\text{min}^{-1}$. Rods which were completely supported, i.e., resting on molybdenum strips, could be heated to 935°C for as long as 10 minutes without change.

HOPKINS and CLARKE[19]* have carried out studies similar to the quenching experiment of MAYBURG and ROTONDI.[16] The surface of a germanium specimen was carefully cleaned and the specimen was sealed in vacuum. By Joulian heating and radiation quenching, the density of acceptors was determined as a function of temperature. The observed acceptor densities were a factor of ten to fifty smaller than those reported by MAYBURG. From this result, HOPKINS and CLARKE concluded that most of the thermal acceptors found by MAYBURG as well as a large fraction of those observed in their own work arose from impurities upon the surface which entered the bulk by diffusion.

A recent experiment performed by LOGAN[20]† indicates that the energy of formation of thermal acceptors is above 2eV. It appears that two acceptor levels, one near the valence band and one in the neighborhood of midband, are introduced into germanium upon quenching from high temperatures. In specimens initially heavily doped with tin ($6\times10^{19}\ \text{cm}^{-3}$), a somewhat higher acceptor density was found. In annealing studies, a monotonic decrease in acceptor density (measured at room temperature) occurred during the first

* The writer is indebted to Dr. CLARKE for the communication of these results prior to publication.

† The author is grateful to Dr. LOGAN for a copy of his paper before publication.

VOL. 1 1956/5

few minutes of annealing. This was followed by a period during which the acceptors were removed very slowly. In a few cases, slight increases in acceptor density were noted in the initial stages of annealing with no further change in the later period.

Table 1. *Data and observations related to thermal conversions in germanium*

Author	Observations with present interpretation
Purdue group[1] and Theuerer and Scaff[1]	Found that *n*-type Ge converted to *p* type upon rapid quenching from high temperatures (probably Cu).
DeSorbo and Dunlap[2]	Thermal acceptor state 0·03 eV above valence band (largely Cu).
Fuller, Theuerer, and Van Roosbroeck[3]	Diffusion coefficient of "thermal acceptors", $D = 0{\cdot}02 \exp(-0{\cdot}52/kT)\text{cm}^2\text{sec}^{-1}$ (Cu).
Taylor[4]	Thermal acceptor density proportional to $\exp[-(1{\cdot}85-5{\cdot}43\times10^{-4}T)/kT]$ (probably vacancies).
Dunlap[5]	Energy of formation of "thermal acceptors" about 1eV (heat of solution of copper).
Fuller and Struthers[6]	Equilibrium density of radio-copper equals equilibrium acceptor density, $6{\cdot}3\times10^{21}\exp(-1{\cdot}2/kT)\ \text{cm}^{-3}$.
Slichter and Kolb[7]	Growing crystals do not convert. Uncut crystals with Cu contaminated surfaces do convert.
Goldberg[8]	Density of thermal acceptors $9{\cdot}4\times10^{22}\exp(-1{\cdot}49/kT)\ \text{cm}^{-3}$ (probably some Cu plus vacancies; plastic deformation).
Esaki[9]	Equilibrium "thermal acceptor" density $2{\cdot}5\times10^{21}\exp(-1{\cdot}2/kT)\ \text{cm}^{-3}$ (Cu). Diffusion coefficient of "thermal acceptors" ($T < 650°\text{C}$), $300\exp(-1{\cdot}6/kT)\ \text{cm}^2\text{sec}^{-1}$ (low-temperature mechanism for Cu).
Morin and Maita[10]	Acceptor level in Cu-doped as well as "pure" Ge 0·040 eV above valence band.
Logan[11]	Average free energy of formation of thermal acceptors 1·8 eV (probably vacancies).
Finn[12]	Cu evaporates quantitatively from vacuum heat-treated Ge.
Seiler *et al.*[13], Maesen *et al.*[14]	Conversion can occur by impurity diffusion from the surface (probably Cu).
Trousil[15]	Vacancies may be annihilated at Ge-Ge melt interface. Cu also may be extracted.
Mayburg and Rotondi[16]	Equilibrium thermal acceptor density is $3\times10^{23}\exp(-2{\cdot}01/kT)\ \text{cm}^{-3}$ (probably vacancies).
Mayburg[17]	Mechanism of annealing of thermal acceptors is complicated; an induction time being observed when density measured at room temperature, a maximum when measured at dry ice temperature. Postulated Frenkel defects (probably vacancies).
Logan and Schwartz[18]	In the absence of plastic deformation, no conversion or change in lifetime is observed if Ge is cooled no faster than $100°\text{C}-\text{min}^{-1}$.
Logan and Schwartz[23]	Cu is extracted from a Ge crystal by surface melts containing Pb or Au.
Logan[36]	Abnormally rapid precipitation of Cu found in Ge with a dislocation density of $10^6\ \text{cm}^{-2}$.
Hopkins and Clarke[19]	Found acceptor densities much lower than those reported by Logan[11, 20] or Mayburg and Rotondi[16] (possibly attributable to sample geometry or to impurity in specimen).
Logan[20]	Energy of formation of thermal acceptors about 2 eV. Density somewhat higher, but formation energy the same in specimens doped with Sn (probably vacancies).
Letaw, Portnoy, and Slifkin[21]	Self-diffusion and vacancy diffusion coefficients in Ge representable by $7{\cdot}8\exp(-2{\cdot}98/kT)\ \text{cm}^2\text{sec}^{-1}$ and $1{\cdot}2\exp(-0{\cdot}96/kT)\ \text{cm}^2\text{sec}^{-1}$, respectively.

2. DISCUSSION

The experiments cited in the previous section led to the conclusion that thermally introduced acceptors (as observed initially) consisted, at least in part, of copper as well as true thermal defects. For convenience, the experimental data cited in the previous section and our interpretation of these data are presented in Table I. The diffusion coefficient of acceptors determined by FULLER, THEUERER, and VAN ROOSBROECK[3] [$D = 0{\cdot}02 \exp(-0{\cdot}52/kT)$ cm^2sec^{-1}] is evidently nearer to that reported for copper by FULLER *et al.*[21] (activation energy $< 0{\cdot}5$ eV) than it is to the diffusion coefficient for a vacancy deduced from self-diffusion experiments,[22] $D_V = 1{\cdot}2 \exp(-0{\cdot}96/kT)$ cm^2sec^{-1}. This expression was obtained by multiplying the self-diffusion coefficient of germanium, $D = 7{\cdot}8 \exp(-2{\cdot}98/kT)$ cm^2sec^{-1},[23] by the reciprocal of the vacancy fraction. The latter quantity was computed by assuming that the acceptor density reported by MAYBURG,[17] $N = 3\times10^{23} \exp(2{\cdot}01/kT)$ cm^{-3}, is a measure of the vacancy density. The acceptor density as a function of temperature at equilibrium was found by GOLDBERG[8] and ESAKI[9] to be $9{\cdot}4\times10^{22}\exp(-1{\cdot}49/kT)$ and $2{\cdot}5\times10^{21}$ $\exp(-1{\cdot}2/kT)$ cm^{-3}, respectively. These densities may be compared to the analogous value determined by FULLER and STRUTHERS[6] for copper, $6{\cdot}3\times10^{21}$ $\exp(-1{\cdot}2/kT)$ cm^{-3}. ESAKI almost certainly observed effects attributable to copper even though the apparent diffusion coefficient for the acceptor was found to be many orders of magnitude lower than that of copper. This disagreement has been ascribed by FULLER *et al.*,[21] either to a surface-limited process or to a difference in mechanism for copper diffusion below 650°C. The limits of error ascertained by GOLDBERG and by FULLER and STRUTHERS cause the activation energies listed above to overlap. Although the departure from linearity in GOLDBERG's data at high temperatures was ascribed by him to cooling before quenching occurs, it is of interest to note that the maximum acceptor density was observed at just that temperature at which the maximum solubility of copper is found. GOLDBERG may also have introduced some thermal defects through plastic deformation in his quenching operation. This mixture of species could account for the observed energy of formation.

Applying the evaporation technique of FINN,[12] MAYBURG[17] found that the density of thermal acceptors as a function of temperature is represented by $N = 3\times10^{23}$ $\exp(-2{\cdot}01/kT)$ cm^{-3}. This appears to be a reliable value in view of the evidence that copper was evaporated from the specimen before the quenching experiments were undertaken and, more important, because of the absence of hysteresis in his quenching series. The fact that an average free energy of formation was reported by LOGAN[11] could readily account for the earlier disagreement with MAYBURG.

The conclusions drawn by TROUSIL[15] from his experiment were that thermal defects are annihilated at an interface between germanium and a melt containing germanium, and that these defects are vacancies. Implicit in his results is the conclusion that thermal defects are not generated in the bulk, for, if they were, reconversion by his technique could not be successful. The presence of the melt, then, requires a large portion of the vacancies introduced in vacuum to be annihilated. One may conclude that, in the absence of plastic deformation, the surface is the only source of thermal acceptors. A possible criticism of this experiment is that under its conditions, copper, or some other acceptor impurity, was more soluble in the melt than in the crystal.[23] Unfortunately, insufficient data were presented with respect to the relative volumes of solid and melt or the cyclic nature of annealing times in the experiment for this question to be resolved here. It is apparent, however, that the thermal equilibrium density of Schottky defects (in the absence of a significant number of sources and sinks such as dislocations) should be controlled by the environment of the crystal. It may be just this effect which accounts for the nonconversion of crystals drawn from a melt of pure germanium.

The annealing experiments of MAYBURG[17] give no direct information as to the region of generation of the thermal imperfections responsible for conversion. The assumption that these are Frenkel defects is not consistent with the data of TROUSIL.[15] At temperatures higher than 516°C it was assumed that annealing occurs by the direct recombination of vacancies and interstitials; however, simple second-order kinetics do not describe the process. Furthermore, even though these defects are considered to be vacancies, the annealing process cannot be represented by the

VOL. 1 1956/57

diffusion of acceptors to the surface. An even more cogent objection to Frenkel defects is that their production requires the highly fortuitous circumstance of identical energies of formation of vacancies and interstitials. It appears, then, that one or the other defect should predominate. This would tend to indicate that thermal acceptors are not Frenkel defects and, if they are Schottky defects, that the annealing process is far from simple.

The Frenkel defect model[17] may be considered unsatisfactory on other grounds. It was proposed that the low temperature mechanism of annealing involves the rapid motion of interstitial germanium ions to dislocations whereupon they are trapped and discharged (not annihilated). Even though each point on a dislocation were suitable for trapping an interstitial, at least 10^7 dislocations/cm^2 would be required to assure that the probability of the collision of an interstitial with a point on a dislocation be at least equal to the probability of collision with a vacant site (using the experimental defect density).[16] This density of dislocations, while abnormally high, still could not account for the behavior observed. Half of the vacancies and interstitials would be expected to be annihilated, reducing the acceptor density to one-half its initial value. In order that the room temperature acceptor density remain constant as observed, highly fortuitous complementary behavior of the rate of trapping of interstitials and the net acceptor gain upon such trappings would have to occur.

HOPKINS and CLARKE[20] carried out a much less exhaustive study than did either MAYBURG[17] or LOGAN;[20] however, their results indicate a consistently lower acceptor density than that found in the latter experiments. The explanation advanced by HOPKINS and CLARKE, e.g., that the surfaces were sufficiently dirty in other experiments to cause doping, does not appear to be the only possible interpretation of their data.* It has been found by use of radioactive Na^{24} that a surface density of the order of 10^{15} sodium ions per square centimeter remained upon germanium surfaces after electrolytic etching in $(Na^{24})_2CO_3$ —NaOH solution followed by (1) three rinses in deionized water at room temperature, (2) a 10-minute dip in deionized water maintained at 65°C, (3) a rinse in chemically pure grade acetone, and (4) drying in a stream of dry nitrogen.[24] This evidence plus the fact that HOPKINS and CLARKE[19] and LOGAN,[20] though using very nearly identical cleaning procedures, obtained different results and MAYBURG and ROTONDI[16] and LOGAN,[20] using different cleaning procedures obtained very similar results, leads one to believe that the surface may not have been involved at all in this contradiction. MAYBURG's suggestion of the presence of some anomalous donor within the material as pulled appear to be more likely. Scaling from the published curve[19] indicates that the density of acceptors introduced in this work may be represented as a function of temperature by $N = 10^{16} \exp(-0{\cdot}6/kT)$ cm^{-3}. It is improbable that either the small pre-exponential factor or the small energy of formation encountered here could be associated with an intrinsic lattice defect having the properties of an acceptor.

The latest experiment of LOGAN[20] agrees well with that of MAYBURG. Using a calculated entropy of formation, LOGAN found that his data were represented as a function of temperature by $N = 1{\cdot}3 \times 10^{24} \exp(-2/kT)$ cm^{-3}, while MAYBURG determined this density to be $N = 3 \times 10^{23} \exp(-2{\cdot}01/kT)$ cm^{-3}.[17] There can hardly be any doubt that the same center was observed by both MAYBURG and LOGAN. The evidence that thermal acceptors are more readily introduced into germanium heavily doped with tin was interpreted by LOGAN to indicate that these acceptors are vacancies. Hall measurements showed that levels were introduced 0·04 eV above the valence band and 0·1–0·2 eV below the conduction band. Though the measurements of acceptor density during annealing were made at room temperature, it was found that a tendency existed for the acceptor density to increase during low-temperature

* Dr. MAYBURG (private communication) has suggested that the small germanium side arms present on the specimen used by CLARKE and HOPKINS[20] may have cooled more rapidly than did the bulk, thereby introducing a sink for thermal acceptors. Dr. CLARKE has observed in previous work [*Phys. Rev.* **95**, 284 (1954)] involving specimens with side arms, a quenched-in acceptor density comparable to that reported by MAYBURG and ROTONDI.[16] It is, however, worthwhile pointing out that in that study, CLARKE encountered a retrograde effect, i.e., a decrease in acceptor density upon heating above 850°C. Such a phenomenon (strongly suggesting the presence of copper) was not encountered by MAYBURG.

anneals. MAYBURG[17] annealed from ten to a hundred times as long as did LOGAN.[20] The initial decrease in acceptors in the latter case may correspond to the presence of a nonintrinsic defect, while the plateau observed agrees completely with room temperature acceptor density determinations made by MAYBURG[17] during the course of the anneal.

LARK-HOROVITZ[25] has proposed that in germanium a vacant site may accept two electrons and that an interstitial may behave as a donor. The presence of both acceptor and donor levels in nucleon and electron bombarded germanium has been confirmed.[26] Fast-neutron bombardment of germanium (CLELAND *et al.*[28]), has been found to introduce levels 0·2 eV below the conduction band and 0·18, 0·07, and 0·02 eV above the valence band. The highest level is presumed to be associated with an interstitial, while the lower two are identified with a vacancy. The origin of the remaining level (0·18 eV above the valence band) is doubtful. Electron bombardment (FAN *et al.*[28]), produced levels 0·22–0·24 eV below the conduction band and 0·11–0·21 eV above the valence band. It has been pointed out that the damage induced by fast-neutron bombardment is much more extensive than that created by 3-MeV electrons.[27] These data do not clearly indicate the number of acceptor or donor levels associated with a vacancy or interstitial, respectively.

It would appear that the data acquired from electron bombardment experiments can be accounted for by a simple model involving only one level associated with each vacancy and interstitial. Multiple levels could arise, for example, from the formation of clusters of vacancies. Doubtless the broadening of levels found in most specimens is evidence of the existence of many regions which are badly damaged and do not consist of a lattice into which widely separated, noninteracting vacancies and interstitials have been introduced.

3. THEORETICAL

3.1 *The Divacancy Model*

As we have pointed out, the annealing process in the high-temperature range is quite complicated. It cannot be described by either a simple first-order or a simple second-order model. We cannot overemphasize the possible error involved in the application to valence crystals of theories established for metals; however, some of these appear to cast light on our problem. SEITZ* has considered the question of the association of vacancies in metals. He concluded that, though vacancies should repel one another at large distances, a short-range attractive force should exist between them for several reasons. The properties of combined pairs of vacancies in copper were calculated by BARTLETT and DIENES,[28] who deduced that the activation energy for the motion of a divacancy in that metal is only 35 to 70 per cent that for the motion of a vacancy. They estimated that the dissociation energy of the vacancy pair in copper is 0·6 eV.

It is worthwhile to consider the formation of divacancies and clusters of higher order as a possible annealing mechanism in thermally converted germanium. It is probable that a combination of processes such as this can describe adequately the observations of annealing above 516°C; however, several apparently anomalous results below that temperature must be accounted for. The first of these is the initial increase in acceptor density with time of annealing when the electrical measurements are carried out at dry-ice temperatures. This is observed even though the density tends to remain constant when measured at room temperature. A second problem is the extremely slow decay of the acceptor density after the maximum is attained.

The increase in acceptor density when measured at −78°C as contrasted to the fact that this does not change when measured at room temperature implies the creation of a level in which electrons are trapped near the Fermi level (perhaps 0·2 eV below the conduction band in this case). This is in good agreement with the Hall data obtained by LOGAN.[20] The acceptor density increases with time and attains a maximum value before final decay. If one assumes that the higher levels are

* F. SEITZ, *Advances in Physics*, (Supplement to *Phil. Mag.*) **1**, 43 (1952). Subsequent to these speculations by SEITZ, some theoretical evidence which indicates that vacancies in metals attract one another has appeared. Papers by J. D. ESHELBY [*Acta Met.* **3**, 487 (1955)] and E. W. MONTROLL and R. B. POTTS [*Phys. Rev.* **100**, 525 (1955)] are of particular importance in this respect. The experimental evidence which is analysed in this section leads us to the conclusion that an activation energy must be provided in order to cause vacancies in germanium to associate. We are unable to resolve this contradiction.

VOL. 1 1956/57

associated with divacancies, the two singly charged vacancies forming this imperfection must, upon association, be able to acquire other electrons. This problem has been studied theoretically by SAXON and HUTNER.[29]* They considered a one-dimensional, monatomic lattice in which two defects are nearest, next-nearest, and second-nearest neighbors. Initially, they showed that in a lattice of this form, the introduction of a single vacancy can result in a new level in the forbidden band. When two vacancies become nearest neighbors, a second level is introduced into the forbidden band. Upon separation, the vacancies begin to behave as individuals, the additional levels vanishing. This analysis, though based upon a highly idealized system, tends to support the assumption of divacancy formation to account for the introduction of additional levels as a result of annealing.

If the second stage of annealing is the diffusion of divacancies and clusters to a point of annihilation, presumably the surface, then the rate of this process is entirely too slow to be in consonance with the theory of BARTLETT and DIENES.[28] It will be recalled that this theory dealt with a metallic lattice (copper). In germanium, a divacancy would be expected to reside on nearest-neighbor or next-nearest-neighbor lattice sites as shown in Fig. 1. In this lattice, there are no common nearest neighbors to a pair of sites. Further, there is no pair of nearest neighbors each of which is nearest neighbor to one of a pair of nearest neighbors. It appears that the most likely mode of motion of a nearest-neighbor divacancy involves the simultaneous jumps of one vacancy to a nearest neighbor and the other vacancy to the site just left by the first vacancy. This amounts to the jump of a single germanium atom past a vacant site and into a next-nearest-neighbor site. This jump distance is a factor of 1·6 times the nearest-neighbor jump distance. From the calculation of SAXON and HUTNER,[29] one sees that it is possible for the divacancy to retain its properties even though separated by an intervening atom; thus, divacancy motion could consist of two discrete vacancy jumps. There is a chance (probably less than one-half[30]) that an atom exchanging sites with a vacancy and residing in the site intermediate to the two vacancies will jump back with a resulting motion of zero for the divacancy. Divacancy motion in a diamond lattice may very well involve a larger activation energy than does simple vacancy motion.

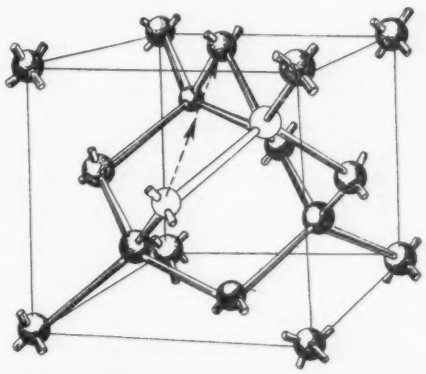


FIG. 1. A nearest-neighbor divacancy is shown on the unshaded sites. Motion of the divacancy cannot be accomplished by simultaneous nearest neighbor jumps of both vacancies to any pair of nearest neighbor, previously unvacated sites. The probable mode of motion, as shown, is the jump of a vacancy to a next-nearest neighbor site.

It is now necessary to investigate the mechanism of formation of a divacancy. The approach which will be used is basically a determination of the collision probabilities of vacancies distributed randomly in the lattice.† The diffusion coefficient of an entity in random motion through a three-dimensional lattice is expressible as $D = d^2\nu/6$, where d is the jump distance and ν is the jump frequency.[31] A diffusing vacancy acquires m new nearest neighbors each time that it jumps and the probability of there being a vacancy on any one of those sites is just N_V/N_S, the fraction of vacancies in the lattice. If N_V vacancies are diffusing, the probability per unit volume and unit time of two vacancies being found on nearest-neighbor sites is $N_V m\nu(N_V/N_S)$. In the diamond cubic lattice, $N_S = 8/a^3$, where a is the lattice constant, $m = 3$, and $d^2 = 3a^2/16$. The collision probability becomes $12N_V{}^2aD_V$, where D_V is the vacancy

* The author is grateful to Dr. KURT LEHOVEC for calling this paper to his attention.

† The author is indebted to Professor JOHN BARDEEN for clarification of this analysis.

diffusion coefficient. If the probability of collision were the same as the probability of divacancy formation, the expression above would describe the rate of the latter process. It appears that an additional term involving the activation energy for the formation of a divacancy must appear. This activation energy may be a measure of the long-range strain interaction of two defects of like sign plus electrostatic terms associated with vacancy repulsion under these conditions. Thus, the probability of a collision which results in the formation of a divacancy is $12N_V^2aD_V \exp(-E_{VVR}/kT)$, where E_{VVR} is the part of the activation energy for the formation of divacancies resulting from the long-range repulsion of vacancies. It is now possible to write

$$-dN_V/dt = 2\alpha N_V^2 \tag{1}$$

where $\alpha = 6aD_V \exp(-E_{VVR}/kT)$ cm^{3}sec^{-1}. The rate of formation of divacancies is just one-half the rate of loss of vacancies, $dN_{VV}/dt = -dN_V/2dt$, where N_{VV} is the divacancy density.

3.2 *Loss of Divacancies by Diffusion*

The process of losing divacancies may be extremely complicated, involving the formation of higher order clusters and the steady annihilation of each of these species at the surface or elsewhere. It will be assumed at first that annealing may be described by the formation of divacancies and their subsequent loss at a surface by means of a diffusion process. This analysis leads to

$$dN_{VV}/dt = \alpha N_V^2 - \beta N_{VV}. \tag{2}$$

The solution of eq. (1) is

$$N_V = N_V^0/(1+2\alpha N_V^0 t), \tag{3}$$

where N_V^0 is the initial vacancy density. By substituting for N_V in eq. (2) the expression eq. (3), we obtain

$$dN_{VV}/dt = \{\alpha(N_V^0)^2/[1+2\alpha N_V^0 t]^2\} - \beta N_{VV}. \tag{4}$$

Using as an integrating factor $\exp(\beta t)$, eq. (4) may be integrated as follows:

$$N_{VV} = \alpha(N_V^0)^2 \exp(-\beta t) \int_0^t \frac{\exp(\beta t)dt}{(1+2\alpha N_V^0 t)^2}. \tag{5}$$

Integration by parts reduces the preceding equation to

$$N_{VV} = \tfrac{1}{2}N_V^0 \exp(-\beta t)\left\{1 - \frac{\exp(\beta t)}{(1+2\alpha N_V^0 t)} + \frac{\beta}{2\alpha N_V^0}\exp\left(-\frac{\beta}{2\alpha N_V^0}\right) I\right\} \tag{6}$$

where, after the change of variable,

$$z = (\beta/2\alpha N_V^0)(1+2\alpha N_V^0 t),$$

$$I = \int_{(\beta/2\alpha N_V^0)}^{(\beta/2\alpha N_V^0)(1+2\alpha N_V^0 t)} \frac{\exp z}{z} dz. \tag{7}$$

The integral in eq. (7) is the exponential integral. It may be integrated, however, by a series expansion which is useful for small values of $(\beta/2\alpha N_V^0)$:

$$I = \ln(1+2\alpha N_V^0 t) + \sum_{k=0}^{\infty}\left\{\left(\frac{\beta}{2\alpha N_V^0}\right)^{k+1}\frac{[(1+2\alpha N_V^0 t)^{k+1}-1]}{K!(K+1)^2}\right\}, \tag{8}$$

and the divacancy density is

$$N_{VV} = \tfrac{1}{2}N_V^0 \exp(-\beta t)\left\{1 - \frac{\exp(\beta t)}{1+2\alpha N_V^0 t} + \left(\frac{\beta}{2\alpha N_V^0}\right)\exp\left(-\frac{\beta}{2\alpha N_V^0}\right)\ln(1+2\alpha N_V^0 t) + \exp\left(-\frac{\beta}{2\alpha N_V^0}\right)\sum_{k=0}^{\infty}\left(\frac{\beta}{2\alpha N_V^0}\right)^{k+2} \times \frac{[(1+2\alpha N_V^0 t)^{k+1}-1]}{K!(K+1)^2}\right\}. \tag{9}$$

The parameters α and β may be evaluated graphically from the data of MAYBURG.[18] This is carried out first writing the expression

$$N_A = N_V + nN_{VV}, \tag{10}$$

where N_A is the number of thermally introduced acceptors and n is the number of electrons associated with a divacancy. Equation (10) may be differentiated with respect to time and, with suitable

VOL. 1 1956/57

substitutions from eqs. (1), (2) and (3), we obtain,

$$dN_A/dt = \frac{\alpha(n-2)(N_V{}^0)^2}{(1+2\alpha N_V{}^0 t)^2} - n\beta N_{VV}. \qquad (11)$$

Evaluating eq. (11) at $t = 0$ or at $t = t_m$ (the time at which the acceptor density is a maximum), we obtain for α and β,

$$\alpha = \frac{[d(N_A/N_V{}^0)/dt]_{t=0}}{(n-2)N_V{}^0},$$

and

$$\beta = \frac{(n-2)\alpha(N_V{}^0)^2}{n(1+2\alpha N_V{}^0 t_m)^2[N_{VV}]_{t_m}}. \qquad (12)$$

One must determine graphically, then, the first slope and the time at which the acceptor density maximizes, $[N_{VV}]_{t_m}$ being computed from eq. (19). The parameters are tabulated in Table 2 where n,

Table 2. Parameters for the divacancy annealing mechanism using first-order annihilation†*

Annealing Temperature (°C)	t_m (sec)	α (cm³sec⁻¹)		β (sec⁻¹)		E_{VVR} (eV)	
		$n = 3$	$n = 4$	$n = 3$	$n = 4$	$n = 3$	$n = 4$
390	$8{\cdot}6\times10^{4}$	$2{\cdot}9\times10^{-20}$	$1{\cdot}45\times10^{-20}$	$1{\cdot}7\times10^{-6}$	$4{\cdot}1\times10^{-6}$	0·77	0·81
402		$4{\cdot}0\times10^{-19}$	$2{\cdot}0\times10^{-19}$			0·66	0·70
447	$6{\cdot}9\times10^{3}$	$1{\cdot}9\times10^{-18}$	$9{\cdot}5\times10^{-19}$	$9{\cdot}5\times10^{-6}$	$2{\cdot}3\times10^{-5}$	0·65	0·69
490	$5{\cdot}2\times10^{2}$	$5{\cdot}7\times10^{-18}$	$2{\cdot}85\times10^{-18}$	$3{\cdot}6\times10^{-4}$	$7{\cdot}6\times10^{-4}$	0·69	0·73

* Experimental data from MAYBURG.[17]

† t_m is the time at which acceptor density maximizes; n is the number of electrons associated with a divacancy at −78°C; E_{VVR} is the activation energy associated with the long-range repulsion of vacancies.

the number of electrons associated with a divacancy at the measuring temperature (−78°C), was chosen to be either three or four. Also shown are the values of E_{VVR} which yield the best representation of α. The long-range repulsive interaction between vacancies in germanium appears to give rise to an activation energy of about 0·7 eV. The experimental [17] and calculated acceptor densities as a function of time are shown in Figs. 2, 3 and 4 for the annealing temperatures 390, 447, and 490°C. The data at the two lower

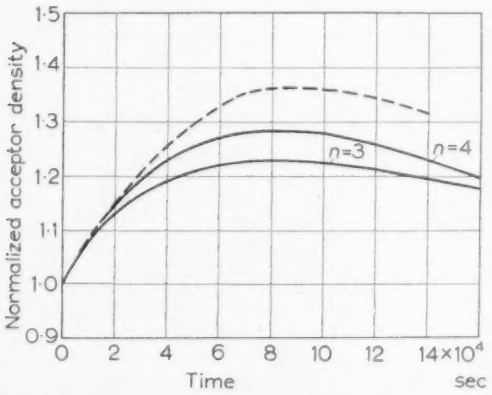


FIG. 2. Experimental (- - -) and theoretical annealing curves at 390°C. Experimental data from MAYBURG.[17] Theory assumes first-order annihilation of divacancies.

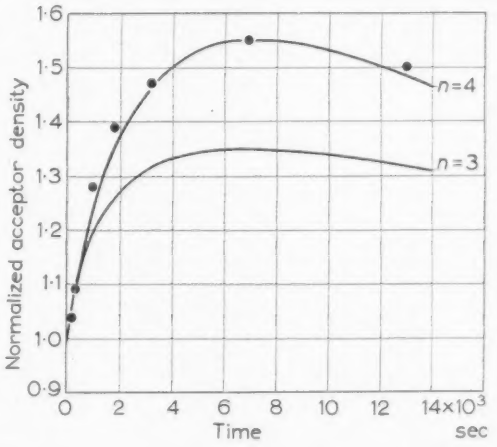


FIG. 3. Experimental points and theoretical annealing curves at 447°C. Experimental data from MAYBURG.[17] Theory assumes first-order annihilation of divacancies.

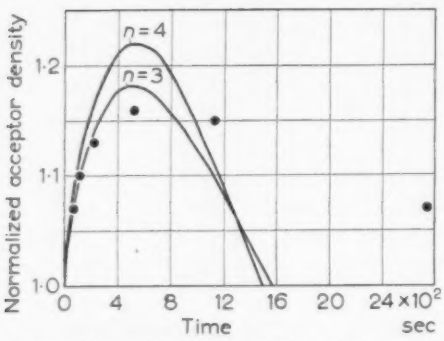


FIG. 4. Experimental points and theoretical annealing curves at 490°C. Experimental data from MAYBURG[17]. Theory assumes first-order annihilation of divacancies.

temperatures are satisfactorily fitted by the theoretical curves; however, it is evident that the process at 490°C cannot be described in this fashion. A ten per cent error in the determination of the first slope can cause a variation of five or so per cent in the maximum acceptor density, but it is impossible to fit the data in question by the assumption of any reasonable error.*

3.3 *Loss of Divacancies by Cluster Formation*

The question of the formation of clusters, i.e., aggregates of more than two vacancies, has been discussed in a previous section. Instead of assuming that a first order, presumably diffusion, process is involved in removing divacancies, let us assume that clusters consisting of four vacancies are formed. Equations (1) and (3) remain unchanged from the previous derivation, but eq. (2) must be modified in order to account for the new decay process, thus,

$$dN_{VV}/dt = \alpha N_V^2 - \beta N_{VV}^2. \qquad (13)$$

By the use of eq. (3) and the substitution $(N_{VV})^{-1} = (1+2\alpha N_V{}^0 t)\varphi$, eq. (13) may be integrated, the divacancy density at any time being

$$N_{VV} = N_V{}^0\left((1+2\alpha N_V{}^0 t)\left\{\left(\frac{\alpha+\beta}{\alpha}\right)^{\frac{1}{2}} \coth\left[\left(\frac{\alpha+\beta}{4\alpha}\right)^{\frac{1}{2}} \ln(1+2\alpha N_V{}^0 t)\right]-1\right\}\right)^{-1}. \qquad (14)$$

We shall assume that in the formation of a cluster of four vacancies, additional electron levels are lost such that only one electron is associated with each of the constituent vacancies, i.e., a total of four electrons with the cluster. One may write, then,

$$N_A = N_V + nN_{VV} + 4N_{4V}, \qquad (15)$$

where N_{4V} is the density of four-clusters. Upon differentiating with respect to time and using the equality $dN_{4V}/dt = -\beta N_{VV}^2/2$, one obtains an expression for α identical to that in eq. (12) and for β,

$$\beta = \frac{\alpha(N_V{}^0)^2}{(1+2\alpha N_V{}^0 t_m)^2[N_{VV}]^2_{t_m}}. \qquad (16)$$

Use of eqs. (12), (14), and (16) and the first slopes and times at which maximum acceptor densities occur permits the parameters to be evaluated. These are shown in Table 3.

The solutions for this case are plotted in Figs. 5, 6, and 7 for the annealing temperatures 390, 447, and 490°C. The criterion of "fit" must lie in the comparison of the reverse slopes of this and the previous case to the experimental data. This treatment, in contrast to the first-order decay,

Table 3. Parameters for the divacancy annealing mechanism using second-order annihilation†‡

Annealing Temperature (°C)	α (cm³sec⁻¹)		β (sec⁻¹)	
	$n=3$	$n=4$	$n=3$	$n=4$
390	$2{\cdot}9\times10^{-20}$	$1{\cdot}45\times10^{-20}$	$7{\cdot}3\times10^{-20}$	$1{\cdot}96\times10^{-19}$
447	$1{\cdot}9\times10^{-18}$	$9{\cdot}5\times10^{-19}$	$5{\cdot}0\times10^{-19}$	$1{\cdot}16\times10^{-18}$
490	$5{\cdot}7\times10^{-18}$	$2{\cdot}85\times10^{-18}$	$3{\cdot}85\times10^{-17}$	$9{\cdot}26\times10^{-17}$

* The author is grateful to Mr. G. H. GOLUB of the staff of the University of Illinois Digital Computer for much assistance in carrying out the integrations which permitted the theoretical curves to be plotted.

† Experimental data from MAYBURG.[17]

‡ n is the number of electrons associated with a divacancy at −78°C.

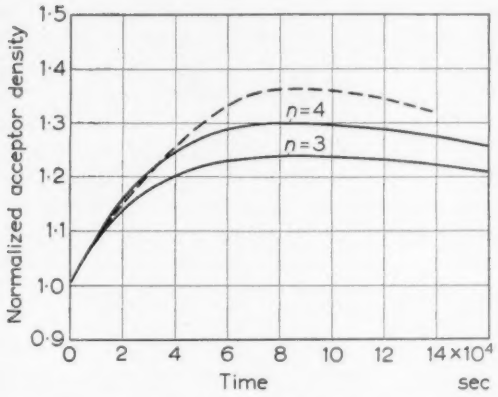


FIG. 5. Experimental (- - -) and theoretical annealing curves at 390°C. Experimental data from MAYBURG.[17] Theory assumes second-order annihilation of divacancies.

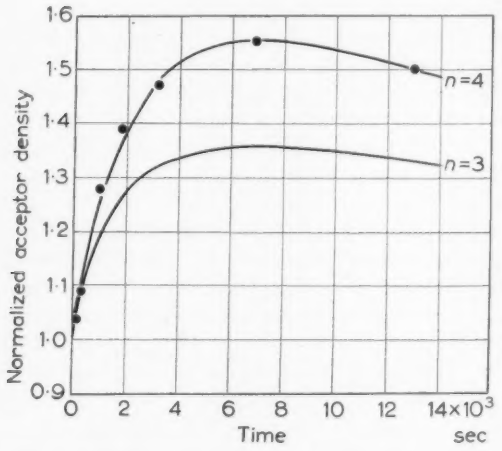


FIG. 6. Experimental points and theoretical annealing curves at 447°C. Experimental data from MAYBURG.[17] Theory assumes second-order annihilation of divacancies.

appears to describe the data well at the highest temperature. Even at the lower temperatures, annealing seems to be somewhat better described by the cluster treatment. An inflection point found in this theoretical decay process is particularly obvious in Fig. 7. This does not correspond to experiment. Equation (15) does not permit the acceptor density to decrease below its initial value since the ultimate conversion of vacancies to clusters of four yields, under our assumptions, no

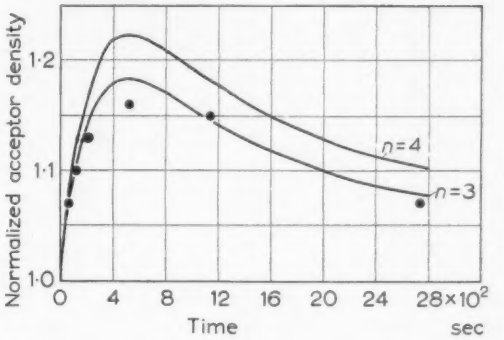


FIG. 7. Experimental points and theoretical annealing curves at 490°C. Experimental data from MAYBURG.[17] Theory assumes second-order annihilation of divacancies.

net change in acceptor density. It is evident that this is not the case; however, this choice was forced upon us because of lack of information as to the nature of the ultimate decay process. The missing factor would, presumably, tend to depress the theoretical curve with the result that the final slope would approach experiment.

3.4 *Properties of Clusters*

The initial assumption that annealing in germanium proceeds through a mechanism involving the formation of divacancies appears to be a satisfactory one. In order to account for the annealing processes observed over the temperature range 390–490°C, two mechanisms for the removal of divacancies have been employed. The first of these is a first-order process as a result of which divacancies are annihilated at a surface, while the second is the formation of clusters of higher order, specifically those containing four vacancies. The latter process appears to be the more important one. Some confirmation of this can be obtained from a presentation of data from a 616°C anneal.[17] Fig. 8 is a replot of these data (scaled from the published curve[17]) arranged so that a second-order process results in a straight line. It appears that two consecutive second-order phenomena are followed by a transition to yet another process about which we have no information. It must be pointed out that insufficient experimental points were taken to establish the first portion of the curve; however, the calculation carried out below provides some justification of this procedure.

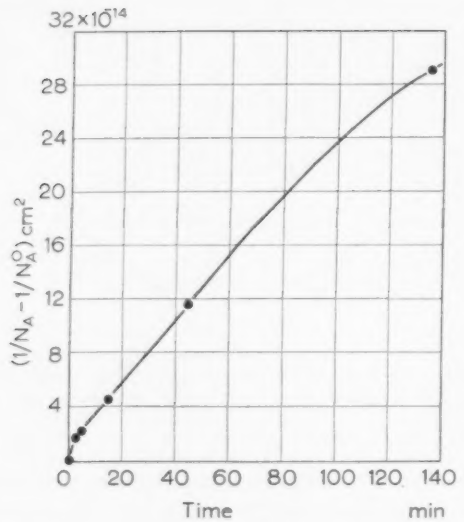


FIG. 8. Experimental annealing data at 616°C plotted so that second-order processes yield straight lines. Experimental data from MAYBURG.[17]

It can be seen from Table 3 that the value of β in the second-order process approaches that of α usually to order of magnitude agreement. To this approximation, we may estimate β at 616°C by use of the expression $\beta \cong \alpha = 6aD_V\exp(-0{\cdot}7/kT)$ $\text{cm}^3\text{sec}^{-1}$, where $D_V = 1{\cdot}2\exp(-0{\cdot}96/kT)\,\text{cm}^2\,\text{sec}^{-1}$.[22] The value thus obtained, $\beta \cong 1{\cdot}7\times10^{-16}$ $\text{cm}^3\text{sec}^{-1}$, may be compared to the second order rate constant determined graphically from the first portion of the curve shown in Fig. 8, $K_1 = 1{\cdot}2\times10^{-16}\,\text{cm}^3\text{sec}^{-1}$. This indicates that the higher temperature process may be the step-wise formation of divacancies, clusters of four and clusters of eight. The rate constant describing the latter step is $K_2 = 4{\cdot}0\times10^{-17}\,\text{cm}^3\text{sec}^{-1}$. The formation of divacancies was not observed, i.e., no maximum was found since it would be expected to occur within less than 20 seconds after the onset of annealing.

Further information on the properties of divacancies in germanium can be extracted from the tabulated values of β for the two cases discussed here. In Section 3.2, we considered the annealing process as being terminated by a diffusion step in which divacancies are annihilated at the surface. If one applies the approximate expression for the diffusion of uniformly distributed imperfections out of a slab, then $\beta = D_{VV}(\pi/W)^2$, where W is the thickness of the slab.[31] Using the data in Table 1 for the case $n = 4$ and substituting for W the smallest dimension of the specimen, 0·1 cm,[17] D_{VV} is found to be represented over this temperature range by the following expression: $D_{VV} = 0{\cdot}1\exp(-2{\cdot}2/kT)\,\text{cm}^2\text{sec}^{-1}$. This result is of doubtful value in view of the better fit obtained by use of the second order decay process, and, in any event, the pre-exponential factor is fixed only within an order of magnitude or so.

The analysis carried out in Section 3.1 for the determination of the rate of formation of divacancies may be applied, with suitable modifications, to the rate of formation of four-clusters. The proportion of the activation energy for four-cluster formation attributable to divacancy repulsion, E_{VR}, may be computed using the values of β in Table 3 and D_{VV} calculated above. One obtains 0·52, 0·61, and 0·56 eV/molecule, respectively, at the three temperatures; thus, E_{VR} is about 0·6 eV. It is more accurate, since four-cluster formation may be dominant at all temperatures, to state simply that the activation energy for the formation of a four-cluster is $D_{VV}+E_{VR} = 2{\cdot}8$ eV, where D_{VV} is the activation energy for divacancy diffusion.

The divacancy mechanism affords insight into an experiment by MAYBURG and ROTONDI[16] referred to in Section 1. It was found that the acceptor density as a function of temperature could be expressed by $N_A = 3\times10^{23}\exp(-2{\cdot}01/kT)$ cm^{-3}. Below 800°, however, the data are fitted quite well by the relationship, $N_A = 1{\cdot}6\times10^{19}$ $\exp(-1{\cdot}1/kT)\,\text{cm}^{-3}$, obtained by scaling from the published graph. The acceptor densities, measured after quenching, were determined at −78°C. Such a decrease in energy of formation can be explained if one assumes that the lower energy is identifiable with the creation of divacancies. Thus we have

$$N_V = 3\times10^{23}\exp(-2{\cdot}0/kT)\,\text{cm}^{-3},$$

and

$$N_{VV} = 4\times10^{18}\exp(-1{\cdot}1/kT)\,\text{cm}^{-3}, \qquad (17)$$

where we have divided the low-temperature acceptor density by four in order to account for the number of electrons we suppose to be associated with a divacancy. We may be certain that we are dealing with equilibrium densities here because

of the linearity of each branch of the semilogarithmic plot from which the relationships were derived. Further, in repeated runs, hysteresis was not observed.

One may write the equilibrium

$$2V \rightleftarrows VV, \tag{18}$$

and the associated equilibrium constant,

$$K_{eq} = N_{VV}/N_V^2 = 4\times10^{-29}\exp(2{\cdot}9/kT)\,\text{cm}^3. \tag{19}$$

The rate of dissociation of divacancies into vacancies can be calculated easily for $K_{eq} = K_F/K_D$, where $K_F = \alpha = 6aD_V\exp(-0{\cdot}7/kT) = 4{\cdot}0\times10^{-7}\exp(-1{\cdot}65/kT)\,\text{cm}^3\text{sec}^{-1}$, the rate of formation of divacancies. Thus, K_D, the rate of dissociation of divacancies, is expressed by

$$K_D = 1\times10^{22}\exp(-4{\cdot}6/kT)\,\text{sec}^{-1}. \tag{20}$$

The extraordinary stability of a divacancy with respect to a vacancy at the lower temperatures is evident from K_{eq}. This may be seen also by a consideration of the energy of formation of two vacancies, 4·0 eV, and the energy of formation of a divacancy, 1·1 eV. The binding energy of a divacancy is, by difference, 2·9 eV.

The Gibbs free energy of any population of vacancies and divacancies can be calculated readily since

$$\Delta G = kT\ln(N_{VV}/K_{eq}N_V^2). \tag{21}$$

The entropy of formation (excluding entropy of mixing) of vacancies and divacancies may be established by solution of the equation

$$N_D = N_S\exp(\Delta S/k)\exp(-\Delta H/kT), \tag{22}$$

through the identification of $N_S\exp(\Delta S/k)$ with the pre-exponential factors of eq. (17). N_D and N_S are, respectively, the density of the defect in question and the density of lattice sites, the latter being $8/a^3$ in the diamond lattice ($4/a^3$ for divacancies). The entropy of formation of a vacancy is $1{\cdot}7\times10^{-4}$ ev-deg^{-1}, while the entropy of formation of a divacancy is $-7{\cdot}4\times10^{-4}$ ev-deg^{-1}. Analyses by VINEYARD and DIENES[32] and by HUNTINGTON, SHIRN, and WAJDA[33] have shown that the introduction of a vacancy should result in an increase in entropy.

4. CONCLUSIONS

The foregoing treatment of the annealing process in germanium, the divacancy model, satisfactorily accounts for the two important anomalies discussed in Section 3.1. The increase in acceptor density in the initial stages of annealing is attributed to the introduction of additional trapping levels for electrons through the formation of divacancies. The slow progress of annealing at low temperatures, despite the rather rapid diffusion of vacancies at these temperatures, is explained either by the small diffusion coefficient for divacancies or, more likely, by the formation of clusters which would also be expected to be relatively immobile.

In this discussion, three-clusters have been omitted from consideration. This was not done on physical grounds, but because of the mathematical difficulties resulting from the inclusion of this species.[34] It is conceivable that the somewhat low calculated maximum density of acceptors at 390°C (Figs. 2 and 5) results from the exclusion of three-clusters. This would follow if, as is probable, more than three electrons were associated with this configuration at dry ice temperatures. As the temperature of annealing was raised, clusters of increasingly higher order appear to have been formed. The experimental data seem to be fitted somewhat better by decreasing n, the number of electrons associated with a divacancy, as the annealing temperature increases. This may be explained on the basis of the thermal instability of divacancies. As we have seen, the rate of dissociation of a divacancy is $K_D = 1\times10^{22}\exp(-4{\cdot}6/kT)\,\text{sec}^{-1}$. This rate is about 43 000 times greater at 490°C than it is at 390°C. The fact that we have omitted this effect from our calculation may have produced the trend noted.

The diffusion of vacancies and clusters plays an important rôle in the processes described. Previous work[22] has shown that the diffusion coefficient of a vacancy in germanium is probably represented by $D_V = 1{\cdot}2\exp(-0{\cdot}96/kT)\,\text{cm}^2\text{sec}^{-1}$. In the present research, a tentative value, $D_{VV} = 0{\cdot}1\exp(-2{\cdot}2/kT)\,\text{cm}^2\text{sec}^{-1}$, has been proposed for divacancy diffusion. Although it is impossible to compare the frequency factors because of the large uncertainty in the latter expression, the activation energies are sufficiently accurate for

comparison. Since the activation energy for diffusion of a divacancy is about a factor of two higher than that of a vacancy, it may be concluded that the analysis of the former process in Sect. 3.1 may be correct. In the diamond lattice, unlike a metal,[28] one would not expect the activation energy for divacancy diffusion to be lower than that for vacancy diffusion. In addition to the points discussed previously, one can arrive at this conclusion from relatively unsophisticated considerations of the nature of bonding in valence crystals as contrasted to bonding in metals.

MAYBURG[25] has calculated the number of jumps of a species with a diffusion coefficient D_V during a radiation quenching operation from 1200 to 300°K. He found that such a defect jumps 2×10^{12} times during cooling. Because of his conclusion that the quenched-in defect is Frenkel type,[17] it was necessary for him to propose that D_V is not the true diffusion coefficient of the thermal defect in question since vacancy-interstitial pairs would be annihilated during the quenching operation. This inference can only lead to the conclusion that diffusion does not occur through a mechanism involving the defects under discussion. Our mechanism provides a more consistent explanation of the observations which have been reported without requiring this stringent condition. Vacancy collisions occurring above 800°C are relatively fruitless insofar as annealing is concerned. The calculated number of jumps (assuming radiation cooling) as the specimen cools through the range below 800°C is only about 1 per cent of the total number of jumps occurring during cooling from 925°C. Over some of this lower range, divacancies are relatively unstable. One may conclude that but a small fraction of the vacancies present in equilibrium at the quenching temperature form divacancies during the quench. The average displacement of a vacancy jumping 10^{12} times in germanium is the order of 0·02 cm. If this resulted in surface annihilation of any but an inconsiderable fraction of the vacancies introduced at the quenching temperature, the excellent linearity of MAYBURG'S[17] plots of logarithm acceptor *versus* reciprocal absolute temperature would be destroyed.

In view of the experimental evidence available at this time, it appears that vacancy formation at high temperatures is responsible for thermal conversion in germanium. Although this effect is important, it is obviously masked by large amounts of copper and other impurities. Such defects as dislocations, if present in large numbers, can readily change the mechanism of annealing from the ones analysed here. The effects of dislocations in germanium on the precipitation of copper have been investigated.[36] It is of interest to note that at 500°C, abnormally rapid precipitation occurs in a germanium specimen containing 10^6 dislocations/cm^2. Upon the completion of his annealing study,[17] MAYBURG found the dislocation density of his specimen to be the order of 10^4–10^5/cm^2.[37]

The best fit to the experimental data appears to be obtained with each divacancy trapping four electrons at −78°C. It is logical to assume that the charge per vacancy decreases with an increase in cluster size if only from the simple argument of the decrease in surface to volume ratio. For this reason, it was assumed that a four-cluster has associated with it four electrons at the temperature in question. The available experimental data did not go over sufficiently long annealing times to make the variation of this parameter worthwhile; thus, we cannot fix with any certainty the net charge of the clusters of four. Furthermore, had the resistivity measurements been made at even lower temperatures, it is possible that the filling of divacancy levels would have been found to be incomplete at dry ice temperatures. One may infer, however, from the Hall data presented by LOGAN that these levels are substantially filled at 200°K.[20]

The activation energy for divacancy formation consists of two parts, E_V, the activation energy for the diffusion of a vacancy, 0·96 eV, and E_{VVR} from the long-range repulsion of two vacancies, 0·7 eV. The activation energy for divacancy formation, then, is about 1·7 eV. The activation energy for the formation of a four-cluster is 2·8 eV. If the activation energy for divacancy diffusion (E_{VV}) is 2·2 eV, as we have estimated, then, the long-range repulsion of two divacancies (E_{4VR}) gives rise to an activation energy of 0·6 eV. This trend appears reasonable in view of the expected increase in activation energy for diffusion with increasing cluster size, and the nature of the strain interactions and electrostatic effects encountered. As the volume of the defect increases, the strain per unit volume of the crystal should decrease.[21]

The experiments of MAYBURG[16, 17] upon

VOL. 1 1956/57

which the present analysis is based are highly consistent internally. The first step in that experiment, the evaporation of acceptors to a density lower than 3×10^{12}cm^{-3}, brought the residual acceptor density down to a value much below the equilibrium solubility of copper at 600°C.[6] If, as a result of improper cleaning, acceptor impurities existed upon the surface, it is evident that they did not enter the specimen during the cleaning period. Furthermore, in view of the complete absence of hysteresis in the large number of heating and quenching cycles carried out by MAYBURG and ROTONDI,[16] it is obvious that the acceptor density was not systematically increased by the diffusion of acceptors from the surface during that period. The recent work of LOGAN[20] is sufficiently corroborative of MAYBURG's[17] that the evaluation of the limited study by HOPKINS and CLARKE[29] becomes very difficult. This problem is more acute in view of the fact that the former studies yielded virtually identical results though employing very different techniques; while the latter experiments, using the cleaning technique of the one[20] and a modification of the heating and quenching technique of the other,[17] agree with neither.

The inference which we have drawn from the various experiments discussed appears to have led to a model which correlates the available data to a satisfactory degree. We have attempted to point out the degree of order existing in the data on thermal conversion and associated phenomena in germanium and to account for these phenomena as well as for annealing and diffusion processes. It is our hope that this study will help to point the way toward further definitive experiments.

5. SUMMARY

1. Thermal acceptors introduced into germanium at high temperatures are lattice vacancies which probably trap one electron at dry ice temperatures and above.

2. The first step in the annealing of excess thermal acceptors in the absence of a high dislocation density is the formation of divacancies by the direct combination of two migrating vacancies. An activation energy of 1·7 eV is associated with this process, 0·96 eV from the activation energy for diffusion of a vacancy and 0·7 eV from the long-range repulsion of two vacancies. Divacancies probably trap two electrons at room temperature and four electrons at dry ice temperatures.

3. The second step of the annealing process may be, at low temperatures, the diffusion of divacancies to the surface, there to be annihilated, or, more probably, the formation of clusters of four vacancies. Four-cluster formation appears to be dominant at intermediate and higher temperatures. The activation energy for the formation of four-clusters is 2·8 eV.

4. Large clusters may be formed at high temperatures.

5. The energies of formation of a vacancy and a divacancy are 2·0 and 1·1 eV respectively. The binding energy of a divacancy is 2·9 eV.

6. The formation of clusters of three was not considered; however, the occurrence of such a step would in no way vitiate the general conclusions.

Acknowledgements.—We gratefully acknowledge the encouragement and assistance of Professor JOHN BARDEEN throughout the course of this research. The members of this Laboratory, particularly Mr. W. M. PORTNOY and Mr. J. R. SCHRIEFFER, provided much illuminating discussion and many helpful suggestions. We are grateful to Professor J. S. KOEHLER for a critical reading of the manuscript and to Dr. S. MACHLUP and Professor R. J. MAURER for several discussions. We wish to express our gratitude to Professor R. E. MEAGHER and the staff of the University of Illinois Digital Computer (ILLIAC) for making many of these computations possible. Discussions with Dr. W. C. DUNLAP, Jr., Dr. R. A. LOGAN, Dr. SUMNER MAYBURG, Professor F. SEITZ, Professor L. SLIFKIN, and Dr. H. STATZ, were of great interest to us.

REFERENCES

1. THEURER H. C. and SCAFF J. H. *Trans. Amer. Inst. Mining Met. Engrs.* **189**; *J. Metals* **4,** 59 (1951). LARK-HOROVITZ K. National Defense Research Council Report NDRC 14–585, March, 1942, to November, 1945.
2. DESORBO W. and DUNLAP Jr. W. C. *Phys. Rev.* **83,** 879 (1951).
3. FULLER C. S., THEUERER H. S. and VAN ROOSBROECK W. *Phys. Rev.* **85,** 678 (1952).
4. TAYLOR W. E. *Phys. Rev.* **86,** 642 (1952).
5. DUNLAP Jr. W. C. *Phys. Rev.* **87,** 190 (1952).

6. FULLER C. S. and STRUTHERS J. D. *Phys. Rev.* **87,** 526 (1952).
7. SLICHTER W. P. and KOLB E. D. *Phys. Rev.* **87,** 527 (1952).
8. GOLDBERG C. *Phys. Rev.* **88,** 920 (1952).
9. ESAKI L. *Phys. Rev.* **89,** 1026 (1952).
10. MORIN F. J. and MAITA J. P. *Phys. Rev.* **90,** 337 (1953).
11. LOGAN R. A. *Phys. Rev.* **91,** 757 (1953).
12. FINN G. *Phys. Rev.* **91,** 754 (1953).
13. SEILER K., GEIST D., KELLER K., and BLANK K. *Naturwissenschaften* **40,** 56 (1953).
14. MAESEN F. V. D., PENNING P., and WIERINGEN A. V., *Philips Research Repts.* **8,** 241 (1953).
15. TROUSIL Z. *Czechoslov. J. Phys.* **4,** 98 (1954).
16. MAYBURG S. and ROTONDI L. *Phys. Rev.* **91,** 1015 (1953).
17. MAYBURG S. *Phys. Rev.* **95,** 38 (1954).
18. LOGAN R. A. and SCHWARTZ M. *Phys. Rev.* **96,** 46 (1954).
19. HOPKINS R. L. and CLARKE E. N. *Phys. Rev.* **100,** 1786 (1955).
20. LOGAN R. A. *Phys. Rev.* **101,** 1455 (1956).
21. FULLER C. S., STRUTHERS J. D., DITZENBERGER J. A., and WOLFSTIRN K. B. *Phys. Rev.* **93,** 1182 (1954).
22. LETAW Jr. H., PORTNOY W. and SLIFKIN L. *Phys. Rev.* **102,** 636 (1956) to be published.
23. LOGAN R. A. and SCHWARTZ M. *J. Appl. Phys.* **26,** 1287 (1955).
24. WOLSKY S., RODRIGUEZ P., and WARING W. Adsorption of Sodium Ions by Germanium Surfaces, Electrochemical Society Meeting, Pittsburgh, 9–19 October, 1955 (*J. Electrochem. Soc.* to be published).
25. LARK-HOROVITZ K. in *Reading Conference on Semiconducting Materials*, 10–15 July, 1950 (Butterworth's Scientific Publications, London, 1951), pp. 47–77.
26. CLELAND J. W., CRAWFORD J. H., and PIGG J. C. *Phys. Rev.* **98,** 1742 (1955); **99,** 1170 (1955); FAN H. Y., KAISER W., KLONTZ E. E., LARK-HOROVITZ K., and PEPPER R. R. *Phys. Rev.* **95,** 1087 (1954).
27. FLETCHER R. C. and BROWN W. L. *Phys. Rev.* **92,** 585 (1953); BROWN W. L., FLETCHER R. C., and WRIGHT K. A. *Phys. Rev.* **92,** 691 (1953).
28. BARTLETT J. H. and DIENES G. J. *Phys. Rev.* **89,** 848 (1953).
29. SAXON D. S. and HUTNER R. A. *Philips Research Repts.* **4,** 81 (1949).
30. BARDEEN J. and HERRING C. *Imperfections in Nearly Perfect Crystals* (John Wiley and Sons, New York, 1952), p. 261.
31. JOST W. *Diffusion in Solids, Liquids, Gases* (Academic Press, Inc., New York, 1952).
32. VINEYARD G. H. and DIENES G. J. *Phys. Rev.* **93,** 265 (1954).
33. HUNTINGTON H. B., SHIRN G. A., and WAJDA E. S. *Phys. Rev.* **99,** 1085 (1955).
34. DIENES G. J. *Phys. Rev.* **91,** 1283 (1953).
35. MAYBURG S. *Phys. Rev.* **98,** 1134 (1955); *Acta Met.* **4,** 52 (1956).
36. LOGAN R. A. *Phys. Rev.* **100,** 615 (1955).
37. MAYBURG S. (private communication). See also *Phys. Rev.* **103,** 1130 (1956).

VOL.
1
1956/57

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 117–127.

PARAMAGNETIC RESONANCE ABSORPTION OF Mn^{++} IN SINGLE CRYSTALS OF ZINCBLENDE

L. M. MATARRESE AND CHIHIRO KIKUCHI*

United States Naval Research Laboratory, Washington, D.C.

(*Received* 7 *June* 1956)

Abstract—The anisotropy of the paramagnetic absorption spectrum of manganous ion in a natural crystal of cubic zinc sulfide has been found to be unusually pronounced. The maximum width of an *hfs* group is 45 gauss with the densities of the components approximately in the ratio 8: 5: 9: 5: 8. This pattern of density distribution is to be expected when quartic terms in the crystalline electric field predominate. The spectra were obtained at fields in the vicinity of 8500 gauss and at a microwave frequency of approximately 24,000 Mc/sec. The angular dependence was investigated by varying the direction of the magnetic field in the (100), (110), and (111) planes of the crystal. Experimental results indicate that the symmetry of the crystalline electric field about the Mn^{++} ion is cubic. The cubic anisotropy constant a is $-1{\cdot}39$ gauss, the sign being determined from the dependence of the *hfs* pattern upon the nuclear magnetic quantum number. The hyperfine interaction constant A is 68·4 gauss. A doublet structure is predicted for the spectrum of manganous ion in hexagonal zinc sulfide.

1. INTRODUCTION

THE usefulness of paramagnetic resonance absorption as a tool in solid state investigations stems from the fact that the localized unpaired electron can be thought of as a microscopic electric and magnetic probe. The latter aspect has been successfully applied to problems in diverse fields, such as the determination of nuclear spins and magnetic moments, F-center studies, donor impurities in silicon, and organic free radicals. The complementary aspect, in which the electron is used to probe the electric fields within crystals, has been applied to such problems as the hexagonal-cubic phase transition in ZnS, nonequivalent sites in calcite, and crystal imperfections in quenched NaCl. The number of applications which have been made in this category is small in comparison to the number of semiconductors, photoconductors, and phosphors to which the method is potentially applicable. Accordingly, it seemed to us worthwhile to investigate some of these potentialities.

A striking example of what can be done along this line is provided in the work of VAN WIERINGEN,[1] who used Mn^{++} ions as indicators to examine the effect of pressure on the transition of ZnS from its hexagonal to its cubic modification. The PMR (paramagnetic resonance) absorption spectrum of small amounts of Mn^{++} in powdered hexagonal ZnS is a complex of 30 or possibly more lines. VAN WIERINGEN showed that as the pressure upon hexagonal ZnS is increased up to about 3500 atmospheres the spectrum gradually changes to that characteristic of Mn^{++} in powdered cubic ZnS, namely one consisting of 6 equally spaced lines. These spectral features had been established earlier[2, 3] in PMR studies of zinc-sulfide phosphors in which the concentration of Mn^{++} was of the order of 0·01 per cent or less. The present contribution seeks to add to our knowledge concerning Mn^{++} in cubic ZnS by an analysis of the PMR absorption spectrum obtained from a single-crystal specimen of zincblende (natural cubic ZnS) in which manganese is fortuitously present as an impurity in a concentration low enough to eliminate unwanted broadening effects. The spectrum we observed is anisotropic and can be explained by assuming that the Mn^{++} ion is located substitutionally in the lattice and is perturbed rather

* Present address: The University of Michigan, Engineering Research Institute, Willow Run Laboratories, Ypsilanti, Michigan, where the concluding phase of this research was carried out.

significantly by a crystalline electric field of cubic symmetry. A similar experiment, involving Mn^{++} in single crystals of calcite, has been reported recently.[4]

2. EXPERIMENTAL PROCEDURE AND RESULTS

The experimental method utilized the 1·2-cm microwave resonance apparatus described by SINGER,[5] the only important modification being to reduce the rate of sweep of the static magnetic field to about 30 gauss/min. The resulting scale calibration on the X–Y recorder was about 18 gauss/in.

The crystal we used for the measurements was a fragment of a large specimen of zincblende of Japanese origin. Another fragment, taken from the same portion

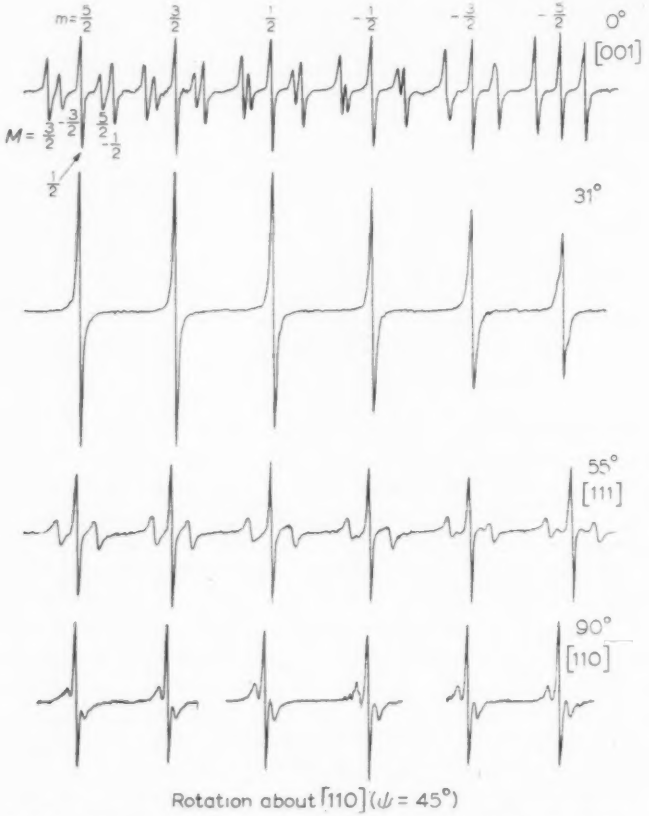


FIG. 1. Experimental records for Mn^{++} in cubic ZnS, at several orientations of magnetic field in the (110) plane. Each reproduced curve is actually a composite of three recorder traces. Magnetic field increases from left to right.

of the parent crystal, was submitted to spectrographic analysis and was reported to contain 0·01–0·1 per cent Mn, an amount which is not inconsistent with the strength of the PMR signal. The other impurities were: Fe, 0·1–1 per cent; Hg, 0·05–0·5 per cent; Al, Co, Cu, Ca, Si, 0·01–0·1 per cent each; Ga, Pb, Mg, 0·001–0·01 per cent each. The mass of the specimen crystal was 26 mg. It was yellow-green in color and had three well-developed cleavages which were identified by X rays as planes of the form {110}. This greatly facilitated the accurate mounting of the crystal in the requisite orientations with respect to the static field direction by the use of optical goniometric techniques. The crystal was affixed with a "tacky" grease to a polystyrene rod protruding from the cavity plunger. In the most important mountings, rotation of the plunger rotated the crystal about its [100], [110], or [111] directions and thereby varied the static field direction in the (100), (110), or

(111) planes, respectively. In this manner the angular dependence of the spectrum was investigated. It is estimated that the overall error in determining the orientation with respect to both the static and the *rf* magnetic field directions was less than $\pm 2°$.

The nonresonant absorption of microwave energy by the crystal was great enough to preclude placing it in the exact center of the cavity. The magnetic field, on the other hand, was measured by a proton resonance coil 3 cm distant from this point. Calibration of the static magnetic field at the position of the crystal was effected by measuring the apparent *g*-value of a sample of hydrazyl placed there. A correction factor (1·0002) was thus obtained, based on a value of 2·0037 for the *g* factor of hydrazyl, and has been applied to all the data quoted here.

Fig. 1 shows the angular dependence of the entire spectrum as the crystal is rotated about a [110] axis. The angle specified is θ, the angle between the [001] axis and the external magnetic field. Since the amplitude of the magnetic field modulation was less than 3 gauss, a value about equal to the width of these lines, the recorder traces very closely approximate the first derivative of the resonance absorption with respect to the magnet current. Thus, the peaks of such absorption lines that are resolved can be located precisely at the points of zero slope, whereas the location of partially resolved lines can only be estimated. In these curves the field increases almost linearly from left to right, from about 8300 to 8700 gauss. The six hyperfine groups are easily identifiable; their spacing increases from 67 gauss at the low-field end to 69 at the high-field end. The center of the spectrum, where the single line of the free ion with zero nuclear spin would appear, has a *g*-value of $2·0025 \pm 0·0002$. This central point and also the positions of the central components of the hyperfine groups do not vary with the direction of the static field by more than the uncertainty involved in determining them—about 0·01 per cent. These data are summarized in Table 1.

The top trace, marked 0°, is obtained when the static field is directed along an edge of the unit cube. Note the five fine-structure components almost completely resolved in the first hyperfine group ($m = 5/2$), with intensities approximately in the ratio 8 : 5 : 9 : 5 : 8. They have here their maximum separation, the distance from the leftmost satellite ($M = 3/2$) peak to the one farthest on the right ($M = -1/2$) being 45 gauss. At the high-field end ($m = -5/2$) the two satellites on either side of the central component ($M = 1/2$) of the group have come together; the separation of the resulting lines is now 38 gauss.

The second orientation in Fig. 1, for which $\theta = 31°$, does not correspond to any simply designated crystallographic direction but is of interest because the fine structure very nearly collapses, especially at the low-field end, yielding a spectrum of six relatively sharp lines.

The remaining spectra in Fig. 1 occur when the applied field is parallel to a body diagonal and a face diagonal, respectively. Note that for $\theta = 55°$ there is a tendency toward greater resolution of the satellites as the field increases, in contradistinction to the case for $\theta = 0°$.

If the crystal is rotated about a [111] axis the

Table 1.

PMR absorption peak positions in gauss for Mn^{++} in cubic ZnS, $H \parallel [001]$, $\nu_0 = 23{,}888$ Mc/sec.
M and *m* are the electronic and nuclear magnetic quantum numbers, respectively. Peak positions are known to $\pm 0·5$ gauss.

M \ *m*	5/2	3/2	1/2	−1/2	−3/2	−5/2
3/2	8325·9	8394·1	8462·2	8531·3		
					8600·3	8670·4
−3/2	8334·9	8400·5	8467·7	8534·5		
1/2	8348·7	8415·6	8483·5	8552·0	8620·9	8690·0
5/2	8362·7	8430·8	8499·8	8569·2		
					8640·0	8709·1
−1/2	8371·0	8437·5	8504·7	8572·2		

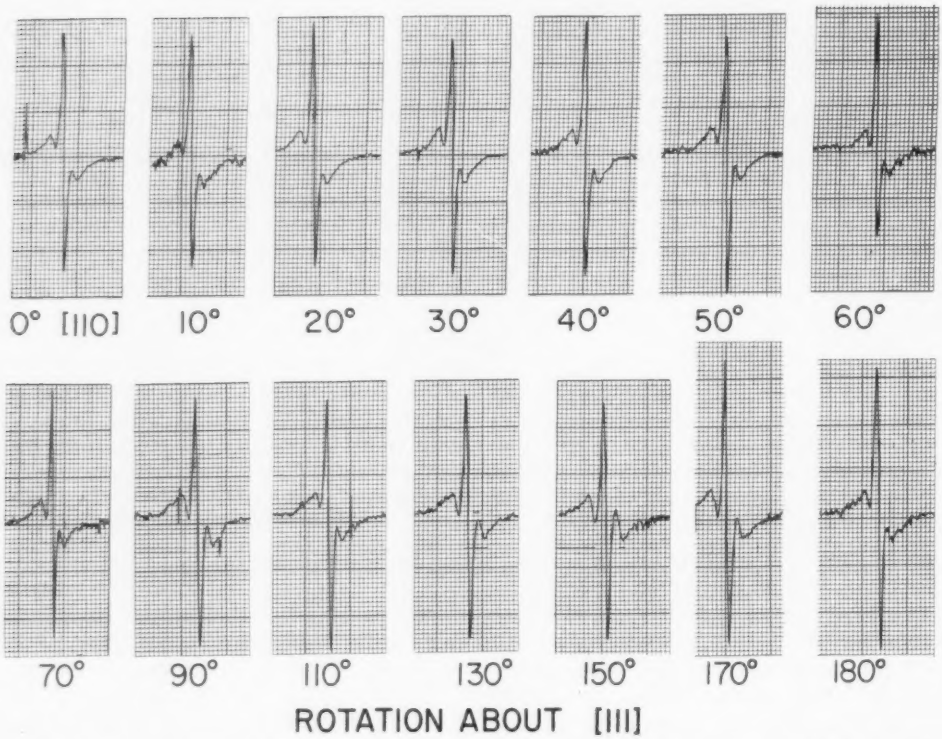


FIG. 2. Experimental records for Mn^{++} in cubic ZnS at several orientations of magnetic field in the (111) plane. Only the first hyperfine group is shown since the others behave similarly. The spikes in the traces for 0°, 90°, and 110° are noise transients.

theory predicts (see below) that the spectrum shall have no angular dependence. This has been experimentally verified. For example, we show in Fig. 2 the behavior of the lowest-field hyperfine group ($m = 5/2$) in this case. Not only is there no shift in position of the group itself, but the fine structure, which is characteristic of that obtained in the [110] direction (see also Fig. 1), does not alter. What differences appear among the traces in the figure can be explained by a slight misalignment of the crystal within the $\pm 2°$ limits and by differences in microwave-power level and noise content.

3. THEORY

A. *Diagonalization of Zeeman Energy*

The interactions relevant to the Mn^{++} problem can be summarized by the perturbation Hamiltonian,[7]

$$\mathcal{H}' = \beta \boldsymbol{S} \cdot \boldsymbol{g} \cdot \boldsymbol{H} + A \boldsymbol{I} \cdot \boldsymbol{S} + V_{\text{cryst}}, \qquad (1)$$

where β is the Bohr magneton, $\boldsymbol{g}$ is a tensor with principal values g_x, g_y, g_z in the crystalline co-ordinate system, $\boldsymbol{S}$ the electron-spin operator, $\boldsymbol{H}$ the applied magnetic field, A is a constant, $\boldsymbol{I}$ is the nuclear spin operator, and V_{cryst} represents the crystalline field potential, which produces the fine structure in the spectrum. The second term gives rise to the hyperfine structure and is expressed as a coupling between the nuclear and electronic magnetic moments of the ion. The first term is the Zeeman term with which we shall be concerned for the moment. Since the Zeeman interaction is about 100 times larger than the hyperfine interaction, which in turn is 10 times as great as the crystalline field interaction, we shall find it convenient to use a co-ordinate system in which the Zeeman term is diagonal. Let $\boldsymbol{H}$ be parallel to the z axis of a co-ordinate system which has the Eulerian angles (θ, ψ, φ) relative to the crystalline

co-ordinate system. Then the problem is to determine expressions for our quantities in a diagonal co-ordinate system that has the Eulerian angles $(\theta', \psi', \varphi')$, not necessarily the same as (θ, ψ, φ) because of the tensor properties of $\boldsymbol{g}$.

The electron spin operator $\boldsymbol{S}$ in the crystal co-ordinate system is related to $\boldsymbol{S}'$ in the diagonal system in the following way:[6]

$$\begin{aligned} S_+ &= \alpha^2 S_+' + 2\alpha\beta^* S_z' - \beta^{*2} S_-', \\ S_- &= -\beta^2 S_+' + 2\alpha^*\beta S_z' + \alpha^{*2} S_-', \\ S_z &= -\alpha\beta S_+' + (\alpha\alpha^* - \beta\beta^*) S_z' - \alpha^*\beta^* S_-', \end{aligned} \tag{2}$$

where $S_\pm = S_x \pm iS_y$ and

$$\alpha = \cos\tfrac{1}{2}\theta\, e^{\frac{1}{2}i(\phi'+\psi')}, \quad \beta = i\sin\tfrac{1}{2}\theta\, e^{\frac{1}{2}i(\phi'-\psi')}. \tag{3}$$

Using $H_x = H\sin\theta\sin\psi$, $H_y = -H\sin\theta\cos\psi$, and $H_z = H\cos\theta$,

the Zeeman term

$$\boldsymbol{S}\cdot\boldsymbol{g}\cdot\boldsymbol{H} = g_x S_x H_x + g_y S_y H_y + g_z S_z H_z \tag{4}$$

becomes

$$\begin{aligned} &S_+'[\tfrac{1}{4}(g_x-g_y)(\alpha^2 H_+ - \beta^2 H_-) + \\ &\quad +\tfrac{1}{4}(g_x+g_y)(\alpha^2 H_- - \beta^2 H_+) - \alpha\beta g_z H_z] + \\ &+S_-'[\tfrac{1}{4}(g_x-g_y)(\alpha^{*2} H_- - \beta^{*2} H_+) + \\ &\quad +\tfrac{1}{4}(g_x+g_y)(\alpha^{*2} H_+ - \beta^{*2} H_-) - \alpha^*\beta^* g_z H_z] + \\ &+S_z'[\tfrac{1}{2}(g_x-g_y)(\alpha\beta^* H_+ + \alpha^*\beta H_-) + \\ &+\tfrac{1}{2}(g_x+g_y)(\alpha\beta^* H_- + \alpha^*\beta H_+) + (\alpha\alpha^* - \beta\beta^*) g_z H_z]. \end{aligned} \tag{5}$$

For this expression to be diagonal, the coefficients of the off-diagonal terms S_+' and S_-' must vanish, i.e.,

$$\begin{aligned} &(g_x-g_y)\sin\theta\sin(\psi'+\psi) - \\ &\quad -(g_x+g_y)\sin\theta\sin(\psi'-\psi) = 0, \end{aligned} \tag{6}$$

$$\begin{aligned} &(g_x-g_y)\sin\theta\cos\theta'\cos(\psi'+\psi) - \\ &\quad -(g_x+g_y)\sin\theta\cos\theta'\cos(\psi'-\psi) + \\ &\quad +2g_z\sin\theta'\cos\theta = 0, \end{aligned} \tag{7}$$

which lead to the solutions

$$\begin{aligned} &\sin\psi' = (g_x/g_\perp)\sin\psi, \quad \cos\psi' = (g_y/g_\perp)\cos\psi, \\ &\sin\theta' = (g_\perp/g)\sin\theta, \quad \cos\theta' = (g_\parallel/g)\cos\theta, \end{aligned} \tag{8}$$

where

$$\begin{aligned} &g_\parallel \equiv g_z, \\ &g_\perp^2 \equiv g_x^2\sin^2\psi + g_y^2\cos^2\psi, \\ &g^2 \equiv g_\parallel\cos^2\theta + g_\perp^2\sin^2\theta. \end{aligned} \tag{9}$$

Under this transformation, the Zeeman term (4) becomes

$$\beta\boldsymbol{S}\cdot\boldsymbol{g}\cdot\boldsymbol{H} = g\beta S_z' H. \tag{10}$$

B. *The Crystalline Electric Field*

The internal electric field with which we are concerned arises from the electrostatic energy of the ions surrounding the paramagnetic ion. It can be shown by means of general arguments,[7] that the crystalline field potential effective in the case of the ions of the iron transition group can be written in the form:

$$V_{\text{cryst}} = V_2 + V_4, \tag{11}$$

where V_2 and V_4 are linear combinations of spherical harmonics of degree 2 and 4 respectively. The harmonics actually appearing in V_2 and V_4 are limited by the crystal symmetry. The cases which are of interest to us are when the crystal z axis is an axis of 3- or 4-fold symmetry. For these cases, the harmonics appearing in V_2 and V_4 are:

	3-fold	4-fold
V_2:	Y_{20}	Y_{20}
V_4:	Y_{40}, $Y_{4\pm3}$	Y_{40}, $Y_{4\pm4}$.

(12)

The particular linear combination that gives a field of proper symmetry is obtained by recognizing that the spherical harmonics Y_{lm} transform like the monomials,[6] $\xi^{l+m}\eta^{l-m}/[(l+m)!(l-m)!]^{\frac{1}{2}}$, where

$$\xi = \alpha\xi' - \beta^*\eta', \quad \eta = \beta\xi' + \alpha^*\eta'. \tag{13}$$

As an example, consider a crystalline field with a z axis of 3-fold symmetry,

$$\begin{aligned} V_4^{(3)} &= b_0 Y_{40} + b_3 Y_{43} + b_{-3} Y_{4-3} \\ &= (b_0/4!)\xi^4\eta^4 + (b_3/\sqrt{7!})\xi^7\eta + (b_{-3}/\sqrt{7!})\xi\eta^7, \end{aligned} \tag{14}$$

where the b's are numerical coefficients. If the crystalline field has cubic symmetry also, the coefficients $b_{\pm3}$ are related to b_0. To see this, let the edge of the cube lie in the xz plane. Now since the crystalline field should be invariant with respect to choice of body diagonal, we can in particular choose a diagonal lying in the xz plane as our new z' axis. There are two of these, corresponding to $\psi' = \pm\frac{1}{2}\pi$. Using (13) and (14) we obtain

$$V_4^{(3)} = (b_0/4!)(\alpha\xi' - \beta^*\eta')^4(\beta\xi' + \alpha^*\eta')^4 + \\ + (b_3/\sqrt{7!})(\alpha\xi' - \beta^*\eta')^7(\beta\xi' + \alpha^*\eta') + \\ + (b_{-3}/\sqrt{7!})(\alpha\xi' - \beta^*\eta')(\beta\xi' + \alpha^*\eta')^7 \\ = (b_0/4!)\xi'^4\eta'^4 + (b_3/\sqrt{7!})\xi'^7\eta' + (b_{-3}/\sqrt{7!})\xi'\eta'^7.$$

Comparing both sides of the second equation, we see that the coefficient of $\xi'^8(Y_{44})$ must vanish, i.e.,

$$(b_0/4!)\alpha^4\beta^4 + (b_3/\sqrt{7!})\alpha^7\beta + (b_{-3}/\sqrt{7!})\alpha\beta^7 \\ = (b_0/4!)\sin^4\theta' + (ib_3/\sqrt{7!})(1+\cos\theta')^3\sin\theta' e^{3i\psi'} - \\ - (ib_{-3}/\sqrt{7!})(1-\cos\theta')^3\sin\theta' e^{-3i\psi'} = 0. \quad (15)$$

For the direction we have chosen we have $\cos\theta' = \frac{1}{3}$ and $\psi' = \pm\frac{1}{2}\pi$; Eq. (15) becomes:

$$(64/4!)b_0 \pm [(64\sqrt{8}/\sqrt{7!})b_3 + (8\sqrt{8}/\sqrt{7!})b_{-3}] = 0. \quad (16)$$

Since the field is real, we must have $b_{-3}Y_{4-3} = (b_3Y_{43})^*$. But $Y_{lm}^* = (-1)^m Y_{l-m}$; therefore $b_{-3} = -b_3^*$. Using this fact, we get from (16)

$$b_3 = \mp(10/7)^{\frac{1}{2}}b_0, \quad (17)$$

which is the desired relation. Thus:

$$V_4^{(3)} = b_0^{(3)}[Y_{40} \mp (10/7)^{\frac{1}{2}}(Y_{43} - Y_{4-3})]. \quad (18)$$

The two crystalline electric fields indicated by the ambiguous sign are related to each other by a rotation of 180° about the body diagonal.

Similarly, for a cubic field about a 4-fold axis, we have

$$V_4^{(4)} = b_0^{(4)}[Y_{40} + (5/14)^{\frac{1}{2}}(Y_{44} + Y_{4-4})] \quad (19)$$

where $b_0^{(4)} = -(3/2)b_0^{(3)}$.

Expressions such as (18) and (19) represent the crystalline electric field in the crystal co-ordinate system. However, as mentioned previously, it is convenient to go over to the diagonal co-ordinate system in order to facilitate the calculation of matrix elements for our problem. In this system, the crystalline field takes the form:

$$V_4 = b_0 \sum_{m=-4}^{4} F_m Y'_{lm},$$

where the coefficients can again be determined by means of the substitutions indicated in (14). For the diagonal elements of (19) we have

$$F_0^{(4)} = 6\alpha^2\alpha^{*2}\beta^2\beta^{*2} - 12\alpha\alpha^*\beta\beta^*(\alpha\alpha^* - \beta\beta^*)^2 + \\ + (\alpha\alpha^* - \beta\beta^*)^4 + \\ + (5/14)^{\frac{1}{2}}(4!/\sqrt{8!})[70(\alpha^{*4}\beta^4 + \alpha^4\beta^{*4})] \\ = \tfrac{1}{8}(35\cos^4\theta' - 30\cos^2\theta' + 3 + 5\sin^4\theta'\cos 4\psi'). \quad (20)$$

Similarly, for the diagonal elements of (18) we have

$$F_0^{(3)} = \tfrac{1}{8}(35\cos^4\theta' - 30\cos^2\theta' + \\ + 3 \mp 20\sqrt{2}\sin^3\theta'\cos\theta'\sin 3\psi'). \quad (21)$$

The primed angles may be eliminated in favor of the unprimed angles (which define the direction of $\boldsymbol{H}$ relative to the crystal axes) by using the relations (8). When $\boldsymbol{g}$ is isotropic, as it turns out to be in our experiment, the distinction between primed and unprimed angles vanishes. Since[4]

$$\langle M|Y'_{40}|M\rangle = \delta[35M^4 - 30J(J+1)M^2 + \\ + 25M^2 - 6J(J+1) + 3J^2(J+1)^2] \\ \equiv \delta f_4(M), \quad (22)$$

where J and M are the electronic spin and magnetic quantum numbers, respectively and δ is a constant whose value can be determined experimentally, we obtain finally

$$\langle M|V_4^{(4)}|M\rangle = (a'/20)F_0^{(4)}(\theta,\psi)f_4(M), \quad (23)$$

$$\langle M|V_4^{(3)}|M\rangle = -(a'/30)F_0^{(3)}(\theta,\psi)f_4(M), \quad (24)$$

in which $18a'$ is the ground state splitting in zero magnetic field.

C. *Hyperfine Structure*

The eigenvalues of the hyperfine-structure Hamiltonian for Mn^{++} are well-known.[8] To second order they are given by:

$$E(M,m) = A'Mm + \\ + \frac{A'^2}{2g\beta H_0}\{M[I(I+1) - m^2] - m[S(S+1) - M^2]\}, \quad (25)$$

where m is the nuclear magnetic quantum number, $g\beta H_0$ is the energy of transition between the unperturbed Zeeman levels of the free ion, and $I = S = J = 5/2$.

D. *Line Positions*

Combining (10), (23), and (25) we obtain the total energy eigenvalue for Mn^{++} in a crystalline electric field of the $V_4^{(4)}$ type. Since the selection rules are $\Delta M = \pm 1$, $\Delta m = 0$, the transition energy is given by $h\nu_0 = E(M,m) - E(M-1,m) = g\beta H(M,m)$ where ν_0 is the (constant) microwave

VOL. 1 1956/57

frequency, and $H(M,m)$ is the resonance magnetic field for the transition. We therefore have

$$H(M,m) = H_0 - Am - \frac{A^2}{2H_0}\left[\frac{35}{4} - m^2 + m(2M-1)\right] + \\ + (a/64)(35\cos^4\theta - 30\cos^2\theta + 3 + 5\sin^4\theta\cos 4\psi)\times \\ \times(56M^3 - 84M^2 - 134M + 81), \quad (26)$$

in units of gauss. In obtaining (26), we substitute $A = A'/g\beta$, $a = a'/g\beta$. The angular dependence of (26) for a single widely-split hyperfine group is shown in Fig. 3 for various values of ψ.

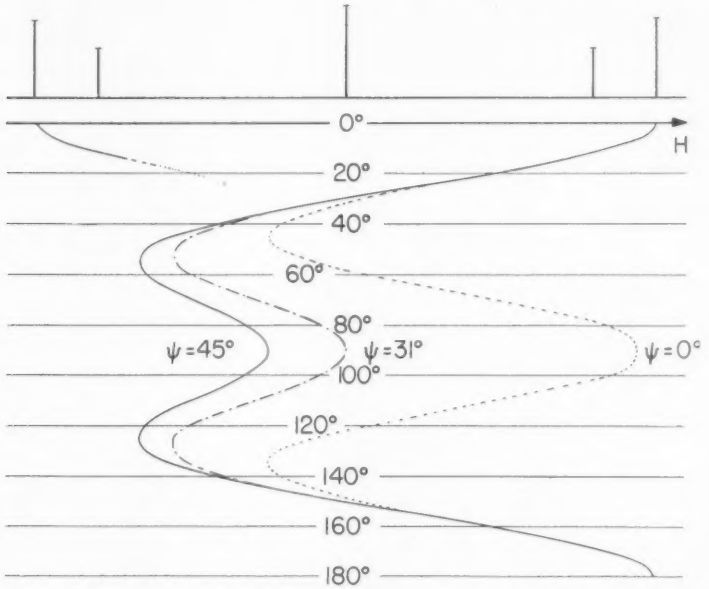


FIG. 3. The angular dependence of a single hyperfine group ($m = 5/2$) for various values of ψ. For clarity only the motion of an outer satellite line is shown.

3. INTERPRETATION

The first step in the interpretation of the spectra is to identify the transitions which yield the observed lines. This can be done readily by noting their relative intensities, which are proportional to $(J+M)(J-M+1)$. These therefore occur in the ratios

$$I_{5/2}:I_{3/2}:I_{1/2}:I_{-1/2}:I_{-3/2} = 5:8:9:8:5, \quad (27)$$

where the subscript denotes the upper of the two magnetic levels involved in a transition and shall hereafter be used to identify the PMR absorption line corresponding to the transition. A further study of the fine structure enables one to tell which of the lines of intensity 5 has the label 5/2 and which of the lines of intensity 8 has the label 3/2. For example, let us particularize (26) for the case in which $\boldsymbol{H}$ rotates in the (110) plane of the crystal and is parallel to [001]. Then $\theta = 0°$, $\psi = 45°$, and

$$H(M,m) = H_0 - Am - \\ -(a/8)(56M^3 - 84M^2 - 134M + 81) - \\ -(A^2/2H_0)[35/4 - m^2 + m(2M-1)], \quad (28)$$

which leads to the fine-structure pattern of Fig. 4 for a given m.

It is of interest to note that the intensity patterns presented by a hyperfine structure group are different for cylindrical (V_2) and cubic (V_4) crystalline electric fields, as indicated by Fig. 5. The observed pattern is consistent with V_4.

The labelling of the hyperfine groups is accomplished by noting that the separation of the central

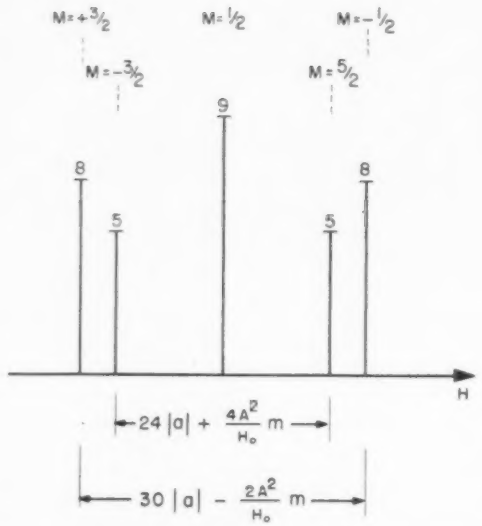


FIG. 4. Schematic representation of fine structure of Mn^{++} spectrum in cubic ZnS for H parallel to [001].

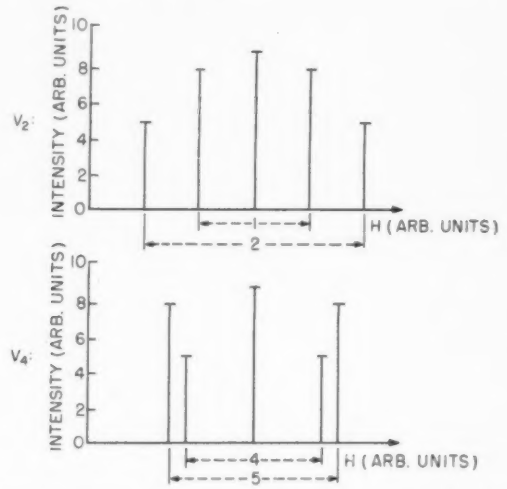
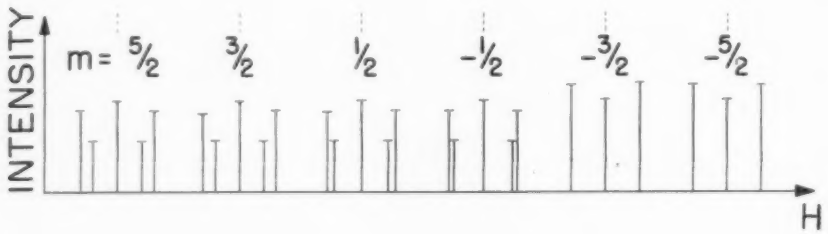


FIG. 5. Intensity pattern of a hyperfine group due to cylindrical (V_2) and cubic (V_4) crystalline electric fields. The strong to weak satellite separation ratios are 1 : 2 for V_2 and 5 : 4 for V_4.



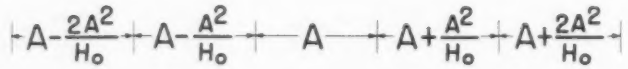


FIG. 6. Schematic representation of entire spectrum of Mn^{++} in cubic ZnS for H parallel to [001], showing separation of the hyperfine groups.

components (intensity 9, $M = \frac{1}{2}$) becomes greater at higher fields. This is shown in Fig. 6, a schematic representation of the spectrum obtained along [001], corresponding to the top trace in Fig. 1.

Having labelled the lines and using the data in Table 1 we may now calculate the parameters in Eq. (28); these are summarized in Table 2. The

Table 2.

Empirical values for the parameters for Mn^{++} *in cubic* ZnS.

H_0	8520·1 ± 0·5 gauss
A	68·4 ± 0·1 gauss
a	−1·39 ± 0·01 gauss
g	2·0025 ± 0·0002

VOL. 1 1956/57

numerical values of A and H_0 are obtained by noting that

$$H(\tfrac{1}{2}, m)-H(\tfrac{1}{2},-m) = -2mA, H(\tfrac{1}{2}, m)+H(\tfrac{1}{2},-m)$$
$$= 2H_0-(A^2/H_0)[(35/4,-m^2)], \qquad (29)$$

from which we obtain $A = 68{\cdot}4$ gauss and $H_0 = 8520{\cdot}1$ gauss.

Experimentally, the positions of the $M = \frac{1}{2}$ lines are independent of orientation of the crystal; hence g is isotropic. Its value calculated from $h\nu_0 = g\beta H_0$, with $\nu_0 = 23{,}888{\cdot}2 \pm 0{\cdot}2$ Mc/s, is $2{\cdot}0025 \pm 0{\cdot}0002$.

From (28) we calculate that for a given hyperfine group the separation of a line of relative intensity 5 and an adjacent line of relative intensity 8 is $3a-3A^2m/H_0$. Thus if a is negative, this separation will be largest for $m = 5/2$ and smallest for $m = -5/2$, which is just the situation experimentally (Fig. 1 and Table 1). The magnitude of a

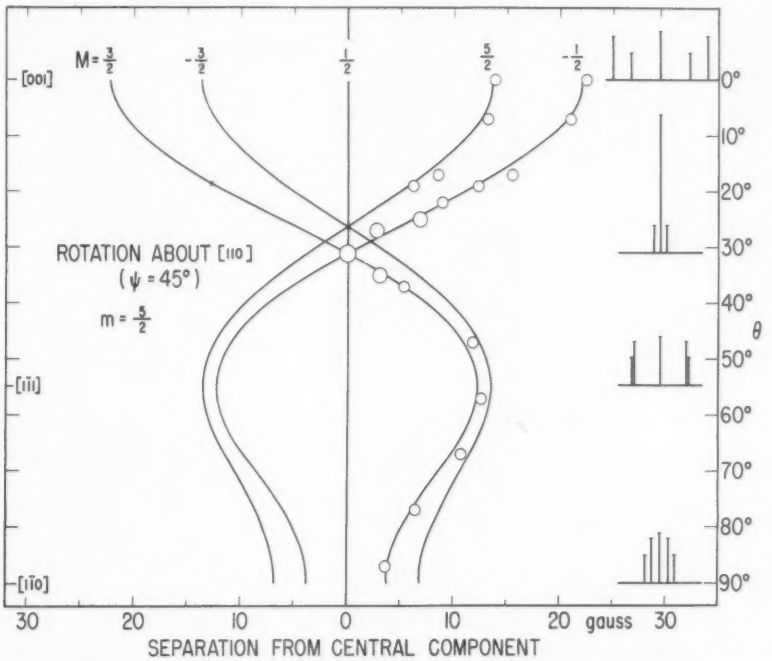


FIG. 7. Comparison of theoretical curves with experimental points (circles) for rotation about [110]. For clarity, only the points on one side are shown.

is best calculated in the obvious manner suggested by Fig. 5. Averaged over the six groups of Table 1, $a = -1{\cdot}39 \pm 0{\cdot}01$ gauss.

The values of these parameters have all been calculated from the observations in a single orientation, namely H parallel to [001]. We now compare the results in other orientations with the predictions of Eq. (26). Two examples are given in Figs. 7 and 8, where the angular dependence of the first hyperfine group over one quadrant is shown for rotation about [110] and [001] respectively. The diagrams at the right of the figure are schematic representations of the spectra at various values of θ. The spectra would look like these diagrams if the lines were infinitely narrow. Compare these patterns in Fig. 7 with the observed spectra in Fig. 1. The experimental points are the positions of peaks in the absorption spectrum. Where the lines are resolved, the peaks fall where expected. Where lines begin to overlap, the peak positions are actually shifted in the direction of the more intense partner, as at $\theta = 35°$, $80°$, or $90°$ in Fig. 7. An interesting feature of Fig. 7 is the manner in which the sharpening-up of the spectrum near $\theta = 31°$

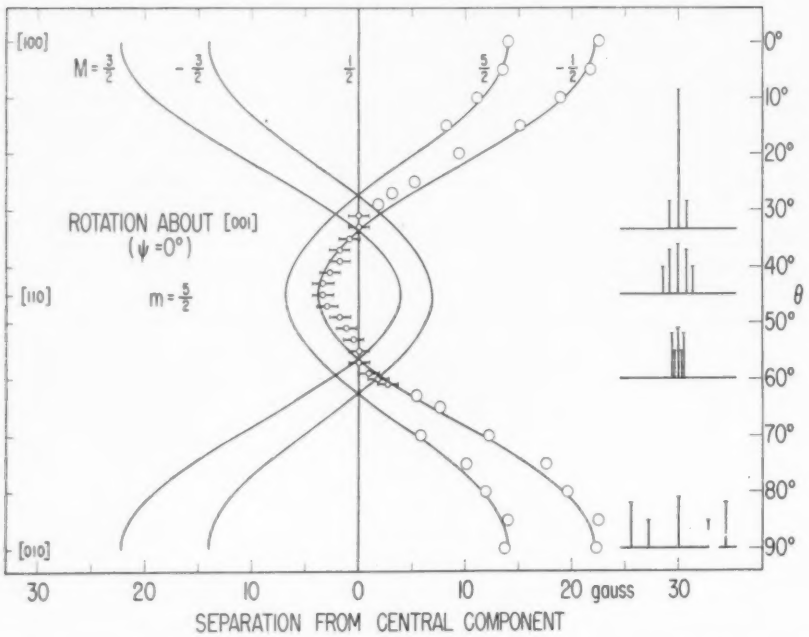


FIG. 8. Comparison of theoretical curves with experimental points (circles) for rotation about [001]. For clarity, only the points on one side are shown.

occurs. Figs. 7 and 8 correspond, respectively, to $\psi = 45°$ and $\psi = 0°$ in Fig. 3. The third curve in this figure, $\psi = 31°$, predicts two sharp-line spectra per period. This also has been verified experimentally.

The situation when the crystal is rotated about [111] is somewhat different. Since this is the 3-fold axis, the angular dependence of the absorption is most easily obtained from (18) and (21). Putting $\theta = 90°$ for the direction of the magnetic field with respect to the axis of rotation, we find that the spectrum is isotropic with a splitting of only one-fourth that found along [100]. This behavior is confirmed by the observations exemplified in Fig. 2.

4. DISCUSSION

Our results indicate that the experimental observations are within the limits of error in agreement with the theoretically expected results for a cubic crystalline electric field. Of particular interest would be a corresponding analysis of Mn^{++} ions in hexagonal ZnS. Because of the difference in the stacking sequence, the cylindrical field V_2 does not vanish, and the absorption spectrum should be characteristic of V_2, perturbed by a cubic field V_4. There are two nonequivalent zinc sites. Consequently, it is anticipated that the absorption spectrum will exhibit the doublet structure such as is found in calcite.[4] An estimate of the doublet splitting can be obtained from (21) and (22), in which the two signs refer to the crystalline electric fields at the two sites respectively. For $M = 3/2$ and $M = 1/2$, the doublet splitting is given by

$$50\sqrt{2}a' \sin^3\theta \cos\theta \sin 3\psi = (75/8)\sqrt{6}\,a',$$

where the last result applies for maximum splitting, occurring at $\psi = 90°$ and $\theta = 60°$. The numerical value of a' in hexagonal ZnS is not known. However, if it is assumed that the nearest neighbors give the major contribution, it can be expected to be close to 1·39 gauss, the value for cubic ZnS. For this value, the calculated splitting is 32 gauss. Stacking disorder, if present, might produce additional structure.

It is possible that ZnS and other related compounds may be useful in the spin measurements of nuclides with small nuclear magnetic moments.

The use of hydrated salts has been exploited by BLEANEY[7] and others. However, in these salts, line broadening of the order of 10 gauss is produced due to the high proton concentration. Such line widths might mask the hyperfine structure completely. By using materials such as ZnS, the undesirable broadening can be reduced. For example, the line width of Mn^{55} in MgO was recently reported to be as small as 1·5 gauss. [9, 10]*

Acknowledgements—We are grateful to Drs. C. M. HERZFELD of the National Bureau of Standards and L. S. SINGER of the National Carbon Research Laboratories for the many stimulating discussions during the early phases of the work presented in this paper.

* They have also reported detecting the signal of V^{51} but the line width is not indicated.

REFERENCES

1. VAN WIERINGEN J. S. *Physica* **19**, 397 (1953).
2. SCHNEIDER E. E. and ENGLAND T. S. *Physica* **17**, 221 (1951).
3. HERSHBERGER W. D. and LEIFER H. N. *Phys. Rev.* **88**, 714 (1952).
4. HURD F. K., SACHS M. and HERSHBERGER W. D. *Phys. Rev.* **93**, 373 (1954); see also KIKUCHI C. *Phys. Rev.* **100**, 1243 (1955).
5. SINGER L. S. *J. Chem. Phys.* **23**, 379 (1955).
6. KRAMERS H. A. *Quantentheorie des Elektrons und der Strahlung*, Edwards Brothers, Lithoprint, 1944, pp. 258, 362.
7. BLEANEY B. and STEVENS K. W. H. *Repts. Prog. Phys.* **16**, 108 (1953).
8. HURD F. K., SACHS M. and HERSHBERGER W. D. *loc. cit.*
9. WERTZ J. E., VIVO J. L. and MUSULIN B. *Phys. Rev.* **100**, 1810 (1955)
10. LOW W. *Phys. Rev.* **101**, 1827 (1956).

BOOK REVIEW

Solid-State Physics. Editors: F. SEITZ and D. TURNBULL. Academic Books Ltd., London. Volume 1, 1955, pp. 469; Volume 2, 1956, pp. 469. 80*s*. each.

SOLID-state physics is a phrase that has been coined since the war. If anybody wants to know what it means, they should look at these two volumes. The subject has its origin in three main pre-war fields of investigation: electron theory of metals and related experimental work on the properties of metals at low temperatures, the Göttingen work on colour centres in alkali halides, and the investigation of dislocations (which started in 1934). Perhaps the most striking development of the postwar years has been the extremely detailed experimental and theoretical investigations of the properties of point defects in germanium and silicon, the introduction of magnetic resonance methods and in particular nuclear resonance as a tool in this field, and experimental observation of dislocations and the proof of their influence on the mechanical, physical, and chemical properties of crystals. One could add the development of neutron diffraction and the study of radiation damage.

These volumes begin with articles by REITZ, by WIGNER and SEITZ, by PINES, and others, on the quantum mechanics of electrons in metals, describing the methods of calculating wave functions and the new considerations introduced by the study of plasma vibrations. There is also an article by FAN summarizing what is known about the valence semiconductors germanium and silicon, a most welcome guide through a vast literature. From Japan, MUTO and TAKAGI write on order-disorder transitions. In the second volume PAKE and KNIGHT have articles on magnetic resonance, SHULL on neutron diffraction, and DE LAUNAY on lattice vibration, and finally there is a detailed report by SEITZ and KOEHLER on displacement of atoms during irradiation. Future volumes promise articles of which there is room to mention only a few—COTTRELL on work-hardening, KITTEL on cyclotron resonance, FRIEDEL on solid-state solutions.

The editors plan to bring out their volumes biannually so as to fulfil within a reasonable time the pressing need for a survey, and thereafter to publish annually. This is the right way to treat a subject like this, which has developed so fast and which impinges on so many other branches of science pure and applied. A review was urgently needed, a review more detailed and more up to date than a series of books could hope to provide. We congratulate the publishers particularly on bringing each volume out when it is ready and in not tying the editors to an annual volume or anything like that; prompt appearance of the reviews will make them all the more useful.

N. F. MOTT

VOL. 1 1956/5

ANNOUNCEMENT

Conference on Magnetism and Magnetic Materials

THE Second Conference on Magnetism and Magnetic Materials has been announced by the American Institute of Electrical Engineers in co-operation with the American Physical Society, the American Institute of Mining and Metallurgical Engineers, and the Institute of Radio Engineers. It will be held on 16, 17, and 18 October at the Hotel Statler, Boston, Massachusetts, U.S.A. This conference, sponsored by the Magnetics Subcommittee of the Basic Science Committee of the American Institute of Electrical Engineers, is designed to bring together individuals interested in basic and applied work on magnetism.

Invited papers will be given by J. H. VAN VLECK; R. M. BOZORTH; T. O. PAINE; C. ZENER and R. R. HEIKES; J. GOODENOUGH; G. H. JONKER, H. P. J. WIJN, and P. B. BRAUN; N. BLOEMBERGEN; D. L. HOGAN; R. L. CONGER; and T. BONN. In addition there will be about 63 contributed papers.

Further information may be obtained by writing to the Local Conference Committee Chairman: T. O. PAINE, Measurements Laboratory, General Electric Company, West Lynn, Massachusetts, U.S.A.

Program

SESSION 1. Tuesday 16 October
8 a.m. Magnetic Anisotropy
SESSION 2. Tuesday 16 October
2 p.m. Permanent Magnets and Fine Particles
SESSION 3a. Wednesday 17 October
9 a.m. Magnetism and Physical Metallurgy
SESSION 3b. Wednesday 17 October
9 a.m. Apparatus and Design
SESSION 4a. Wednesday 17 October
2 p.m. Losses in Soft Magnetic Materials
SESSION 4b. Wednesday 17 October
2 p.m. Ferrites
SESSION 5. Thursday 18 October
9 a.m. High-Frequency Phenomena
SESSION 6. Thursday 18 October
2 p.m. Switching Devices and Magnetic Amplifiers

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 129–136.

FERROMAGNETIC RESONANCE LINE WIDTH IN INSULATING MATERIALS

A. M. CLOGSTON, H. SUHL, L. R. WALKER and P. W. ANDERSON

Bell Telephone Laboratories, Murray Hill, New Jersey

(*Received* 2 *July* 1956)

Abstract—It has recently been recognized that the energy spectrum of spin waves in a ferromagnetic body of finite size is very different from the spectrum in an infinitely extended body. In a finite sample a degeneracy exists between the spin wave of wave vector $\kappa = 0$ and a manifold of short wavelength spin waves. In the presence of irregularities in the magnetic lattice, such as are necessarily present in magnetic insulators having an inverted spinel structure, the degeneracy leads to a resonance line width comparable to those experimentally observed. This line width depends upon the magnetization of the material, the shape of the sample, the Curie temperature and the degree of ordering of the magnetic lattice.

INTRODUCTION

A CONSIDERABLE body of experimental measurements has accumulated in the last few years dealing with gyromagnetic resonance in highly insulating ferromagnetic materials such as the various ferrites.[1, 2, 3] There are two important comments that can be made about the resonance line widths observed in these experiments. First, the narrowest lines so far measured on the best materials under ideal conditions lie in the neighborhood of 20 to 50 oersteds. Secondly, the line widths tend to increase as the temperature is lowered, suggesting strongly that a finite line width would remain even at absolute zero.

In some particular instances a definite mechanism has been identified as contributing largely or in part to the observed line width. One such case is the valence exchange effect encountered by GALT[3] in iron-rich nickel iron ferrite. For metallic ferromagnets another process has been described by AMENT and RADO[4] where the spin system is coupled to the electromagnetic field by a skin depth effect. In these two cases the precessional motion of the spins is coupled in a particular way to the translational motion of the electrons and the spin energy is dissipated in the electronic degrees of freedom.

In highly insulating materials, however, such as stoichiometric manganese ferrite, neither of these mechanisms can be used to explain the relaxation process. The mystery is deepened by the fact that the direct interaction between the spin system, in a mode of uniform precession, and the lattice is very small. The strength of this interaction is related to the size of the magnetostrictive coefficients and is generally agreed to lead to a relaxation time of order of 10^{-1} seconds or longer.[5] A line width of 50 oersteds, on the other hand, corresponds to a relaxation time of about 10^{-8} seconds. There appears to be, therefore, in these insulating ferromagnets, a source of line width as yet unidentified.

Several attempts have been made to examine the effect of dipolar forces on an array of similar spins, situated at the points of a regular lattice, and interacting mainly by means of an exchange energy, in the hope of demonstrating the existence of an intrinsic resonance line width. The first such attempt was made by VANVLECK[6] using the method of moments. VanVleck's calculation was based on a truncated Hamiltonian and predicted a strongly temperature dependent line width going to zero at absolute zero of temperature. VanVleck truncated his Hamiltonian as a matter of principle to avoid including in the moment calculation the effect of satellite lines at frequencies remote from the main resonance. The strong temperature dependence of his result makes it impossible to account in this way for the experimental behavior.

KEFFER,[7] KITTEL and ABRAHAMS[5] have criticized VanVleck's calculation. They argue that, in the presence of strong exchange interaction, the satellite terms excluded by VanVleck actually contribute to the broadening of the main resonance and should be included in a proper moment calculation. Proceeding on this basis, Keffer worked out the second moment at absolute zero and found a non-vanishing result. It was contended by him and by Kittel and Abrahams that an explanation had been obtained of the existence of a low temperature line width with roughly the observed order of magnitude.

Keffer's position has since been challenged by BLOEMBERGEN and WANG[8] who maintain that the terms leading to the low temperature second moment cannot contribute to the broadening of the main line. We have examined this question in detail from the point of view of spin wave theory and concur with Bloembergen and Wang. This same result is also implicit in the work of KASUYA.[9] It appears therefore, that it is not possible to account for a line width of the observed size persisting to absolute zero, on the basis of dipolar forces alone.

In looking further afield for a source of ferromagnetic resonance line width, it is natural to consider the effect of magnetic inhomogeneities in the experimental materials. This question was raised by VANVLECK[6] and by BLOEMBERGEN and WANG.[8] In the case of the more or less inverted spinels, such as nickel or manganese ferrite, the point becomes particularly relevant. The well known random distribution of two kinds of magnetic ions on the octahedral sites of the lattice will certainly lead to a random variation of the pseudo-dipolar forces between neighboring ions. The variation will be of the order of the pseudo-dipolar field and amount to perhaps 10^5 oersted, a magnitude hardly to be overlooked.

In an early calculation of the effect of this random variation one of us arrived at a result predicting a finite but rather small line width. This result was based on a spin wave spectrum appropriate to an infinitely extended magnetic sample, such as found by HOLSTEIN and PRIMAKOFF[10] and KITTEL and HERRING.[11] In a study of ferromagnetic resonance at high signal power levels by ANDERSON and SUHL,[12] it became apparent that the spin wave spectrum must be drastically modified when considering a finite magnetic body. The appropriate spectrum for an ellipsoid of revolution magnetized along the symmetry axis is given by

$$\omega_\kappa = \gamma\sqrt{(H-N_zM+H_ea^2\kappa^2)(H-N_zM+H_ea^2\kappa^2+4\pi M\sin^2\theta)} \tag{1}$$

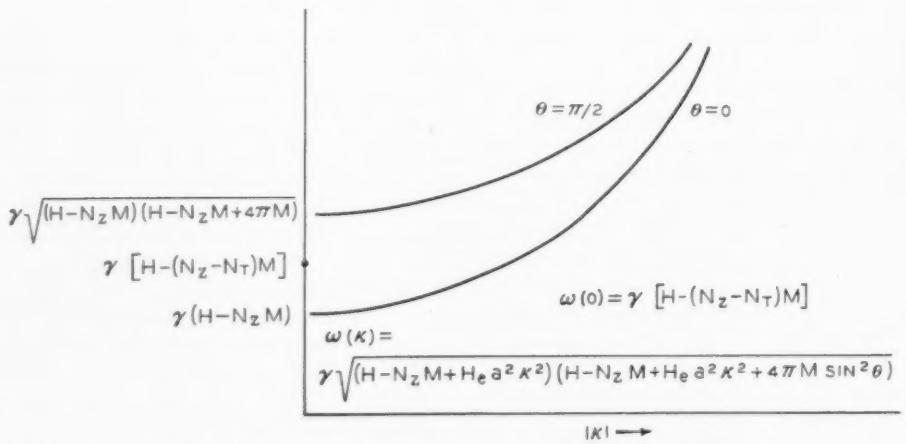


FIG. 1. Dispersion Relation for a Finite Body of Ellipsoidal Shape.

and is shown in Fig. 1. The energy of a spin wave with wave vector κ and an excitation of one quantum is given by $\hbar\omega_\kappa$, where ω_κ depends upon the magnitude of κ and the angle θ made by κ with the axis of quantization. If this direction is taken to be the z-axis, then N_z is the corresponding demagnetization factor and N_t is the demagnetization factor in all directions perpendicular to the z-axis. M is the saturation magnetization of the material, a is the nearest neighbor distance and H_e is an exchange field. The quantity γ is equal to $g\beta/\hbar$, where g is the gyromagnetic ratio and β the Bohr magneton. Eq. (1) holds for $a\kappa \ll 2\pi$ and would require modification near the edges of the Brillouin zone. The frequency of the "spin-wave", $\kappa = 0$, corresponding to a uniform precession of the spin, which is the one usually excited directly in a resonance experiment satisfies a different relation. It is given by

$$\omega_0 = \gamma[H-(N_z-N_t)M], \tag{2}$$

the well-known Kittel relation when anisotropy is neglected.[13] For $\kappa > 0$ but very small so that the corresponding wavelength is of the order of the sample size, spin waves are no longer even approximately good wave functions and must be replaced by a more appropriate representation. This question will be the subject of a forthcoming publication by one of us (L.R.W.) and will not be considered further here.

The outstanding feature of the spin wave spectrum shown in Fig. 1 is the degeneracy of ω_0 with many spin waves of large κ value. In an infinite medium with $N_z = 0$, $N_t = 0$, ω_0 lies at the very bottom of the spectrum and leads to the unacceptably narrow line width first calculated from the random variations of the pseudo-dipolar fields. In the presence of the degeneracy, however, a line width is obtained that can be easily of the observed order of magnitude.

Eqs. (1) and (2) which express the spin wave degeneracy will be a result of the theory given in the second section of this paper. It is easy enough, however, to see the source of the degeneracy from a physical point of view. For a spin wave of large κ number, the transverse internal field produced by poles appearing at the surface of the sample will be averaged out, except very near the surface, whereas the longitudinal field will be reduced by the demagnetizing field N_zM. For the $\kappa = 0$ spin wave, however, the transverse field is not averaged out and tends to compensate the effect of the longitudinal demagnetization. Consequently, the short wavelength spectrum is depressed from the infinite case by a greater amount than the spin wave for $\kappa = 0$ and the degeneracy follows.

2. THEORY

The theory that we are now to discuss essentially describes the scattering of spin waves by disorder of any sort in a regular lattice of coupled spins. In order to make progress with this theory of disorder scattering, we adopt a greatly abstracted model, consisting of an array of identical spins interacting by means of a ferromagnetic exchange integral as well as ordinary dipolar and pseudo-dipolar terms. We thus ignore any conduction electron effects and also set to one side any sub-lattice effects that might arise from the ferromagnetic structure of the various magnetic insulators. We adopt, therefore, the following Hamiltonian for the system,

$$\mathcal{H} = -g\beta H\sum_j S_{jz} - \sum_{jk} \boldsymbol{J}_{jk}\boldsymbol{S}_j\cdot\boldsymbol{S}_k + \\ +\frac{1}{2}\sum_{jk} D_{jk}\left[\boldsymbol{S}_j\cdot\boldsymbol{S}_k - \frac{3}{r_{jk}^2}(\boldsymbol{r}_{jk}\cdot\boldsymbol{S}_j)(\boldsymbol{r}_{jk}\cdot\boldsymbol{S}_k)\right] \tag{3}$$

where D_{jk} includes both ordinary and pseudo-dipolar components. We now express the Hamiltonian in terms of the variables $S_{j+} = S_{jx}+iS_{jy}$ and $S_{j-} = S_{jx}-iS_{jy}$ and further introduce the abbreviations

$$A_{jk} = -2\boldsymbol{J}_{jk}+\tfrac{1}{2}D_{jk}(3\gamma_{jk}^2-1) \tag{4}$$

$$B_{jk} = -\tfrac{3}{2}D_{jk}(3\gamma_{jk}^2-1) \tag{5}$$

$$E_{jk} = -\tfrac{3}{4}D_{jk}(\alpha_{jk}-i\beta_{jk})^2 \tag{6}$$

$$F_{jk} = -\tfrac{3}{2}D_{jk}\gamma_{jk}(\alpha_{jk}-i\beta_{jk}) \tag{7}$$

obtaining thereby

$$\mathcal{H} = -Hg\beta\,\Sigma S_{j_z}+\tfrac{1}{2}\sum_{jk}A_{jk}\boldsymbol{S}_j\cdot\boldsymbol{S}_k+\tfrac{1}{2}\sum_{jk}B_{jk}S_{jz}S_{kz} \\ +\tfrac{1}{2}\sum_{jk}(E_{jk}S_{j+}S_{k+}+E_{jk}{}^*S_{j-}S_{k-})+ \\ +\Sigma(F_{jk}S_{j+}S_{kz}+F_{jk}{}^*S_{j-}S_{kz}) \tag{8}$$

The quantities α_{jk}, β_{jk} and γ_{jk} are direction cosines of the vector $\boldsymbol{r}_{jk}$ joining spins j and k. The notation adopted here is the same as that used by KEFFER.[7]

We now form the commutator between $\mathcal{H}$ and S_{j+} to obtain the equations of motion

$$i\hbar\dot{S}_{j+} = Hg\beta S_{j+} + \sum_k \{A_{jk}(S_{jz}S_{k+} - S_{j+}S_{kz}) - B_{jk}S_{j+}S_{kz} + 2E_{jk}{}^*S_{jz}S_{k-} - F_{jk}S_{j+}S_{k+} - F_{jk}{}^*S_{j+}S_{k-} + 2F_{jk}{}^*S_{jz}S_{kz}\} \quad (9)$$

These equations are next linearized by replacing S_{jz} everywhere by its expectation value, S, and dropping terms in $S_{j+}S_{k+}$ or $S_{j+}S_{k-}$. We have then

$$i\hbar\dot{S}_{j+} = Hg\beta S_{j+} + S\sum_k \{A_{jk}(S_{k+} - S_{j+}) - B_{jk}S_{j+} + 2E_{jk}{}^*S_{k-}\} \quad (10)$$

It is assumed here that the symmetry of the body is great enough, and the disorder in the lattice random enough, so that $\sum_k F_{jk}{}^*$ is zero. We require the physical situation, in other words, to be such that S_{j+} has no constant contribution. Let us next expand S_{j+} in a Fourier series

$$S_{j+} = \sqrt{\frac{2S}{N}} \sum_\kappa e^{i\kappa\cdot r_j}\sigma_\kappa \quad (11)$$

thereby introducing the new set of dynamical variables σ_κ. It should be noted particularly that the total spin $S_+ = \Sigma S_{j+}$ is given by $\sqrt{2SN}\,\sigma_0$. The equation of motion for σ_κ is found to be

$$i\hbar\dot{\sigma}_\kappa = Hg\beta\sigma_\kappa + \frac{S}{N}\sum_{\kappa'}\sigma_{\kappa'}\sum_j e^{i(\kappa'-\kappa)\cdot r_j}\times \times \sum_k \{A_{jk}[e^{i\kappa'\cdot(r_k-r_j)}-1] - B_{jk}\} + \frac{2S}{N}\sum_{\kappa'}\sigma^*_{-\kappa'}\sum_j e^{i(\kappa'-\kappa)\cdot r_j}\times \times \sum_k E_{jk}{}^* e^{i\kappa'\cdot(r_k-r_j)} \quad (12)$$

The variable $\sigma_\kappa{}^*$ is the Hermitian conjugate of σ_κ. We shall now explicitly recognize the disorder of the lattice by placing $A_{jk} = A_{jk}{}^\circ + a_{jk}$, $B_{jk} = B_{jk}{}^\circ + b_{jk}$ and $E_{jk} = E_{jk}{}^\circ + e_{jk}$. The quantities $A_{jk}{}^\circ$, $B_{jk}{}^\circ$ and $E_{jk}{}^\circ$ will be related to an average exchange integral $J_{jk}{}^\circ$ and average dipole coefficient $D_{jk}{}^\circ$ by relations identical to equations (4) to (7), and will be uniform throughout the lattice. Equation (12) becomes

$$i\hbar\dot{\sigma}_\kappa = Hg\beta\sigma_\kappa + \frac{S}{N}\sum_{\kappa'}\sigma_{\kappa'}\sum_j e^{i(\kappa'-\kappa)\cdot r_j}\times \times \sum_k \{A_{jk}{}^\circ[e^{i\kappa'\cdot(r_k-r_j}-1] - B_{jk}{}^\circ\} + \frac{2S}{N}\sum_{\kappa'}\sigma_{-\kappa'}{}^*\times \times \sum_j e^{i(\kappa'-\kappa)\cdot r_j}\sum_k E_{jk}{}^{\circ *} e^{i\kappa'\cdot(r_k-r_j)} + \frac{S}{N}\sum_{\kappa'}\sigma_{\kappa'}\times \times \sum_j e^{i(\kappa'-\kappa)\cdot r_j}\sum_k \{a_{jk}[e^{i\kappa'\cdot(r_k-r_j)}-1] - b_{jk}\} + \frac{2S}{N}\sum_{\kappa'}\sigma_{-\kappa'}{}^*\sum_j e^{i(\kappa'-\kappa)\cdot r_j}\sum_k e_{jk}{}^* e^{i\kappa'\cdot(r_k-r_j)} \quad (13)$$

We will now assume that the various sums in eq. (13) involving $A_{jk}{}^\circ$, $B_{jk}{}^\circ$ and $E_{jk}{}^\circ$ taken over the index k are independent of the point of origin j except possibly in a very small region near the surface of the body. If $\kappa' = 0$ this will be true for any ellipsoid, and will be true for any shape sample if κ' corresponds to a wavelength short compared to the sample dimensions. The assumption will not be true for long spin waves; the resulting error in the calculations should not be important unless the disorder of the lattice is such as to particularly emphasize the role of these long spin waves in the relaxation process.

The sums over $A_{jk}{}^\circ$, $B_{jk}{}^\circ$ and $E_{jk}{}^\circ$ may be explicitly evaluated* for $\kappa' = 0$, or κ' large, and lead to somewhat different equations of motion for σ_0 and σ_κ as follows.

$$\dot{\sigma}_0 = i\gamma[H - M(N_z - N_t)]\sigma_0 - \frac{1}{i\hbar}\frac{S}{N}\sum_k \sigma_\kappa \times \times \sum_{jk} e^{i\kappa\cdot r_j}\, b_{jk} + \frac{1}{i\hbar}\frac{2S}{N}\sum_\kappa \sigma_{-\kappa}{}^*\sum_{jk} e^{i\kappa\cdot r_j}\, e_{jk}{}^* \quad (14)$$

$$\dot{\sigma}_\kappa = i\gamma[H - MN_z + H_e a^2\kappa^2 + 2\pi M\sin^2\theta_\kappa]\sigma_\kappa + i\gamma[2\pi M\sin^2\theta_\kappa e^{2i\phi_\kappa}]\sigma_{-\kappa}{}^* - \frac{1}{i\hbar}\frac{S}{N}\sigma_0\sum_k e^{-i\kappa\cdot r_j} b_{jk} + \frac{1}{i\hbar}\frac{2S}{N}\sigma_0{}^*\sum_{jk} e^{-i\kappa\cdot r_j} e_{jk}{}^* \quad (15)$$

Here N_z and N_t are the demagnetizing factors of the sample $\gamma = -g\beta/\hbar$, $H_e = \frac{1}{3}JZS$, $\cos\theta_\kappa = \kappa_z/\kappa$ and $\tan\varphi_\kappa = \kappa_y/\kappa_x$. In the equation for $\dot{\sigma}_\kappa$ we have ignored interactions produced by the disorder with all spin waves except $\kappa = 0$, on the grounds that in the problem we shall be considering σ_0 will be much larger than any other σ_κ. It may be noted that fluctuations in the exchange field do not contribute to the interaction terms in equations (14) and (15).

We shall now make certain changes of variable first introduced by HOLSTEIN and PRIMAKOFF,[10] as follows

$$\sigma_\kappa = \cosh\frac{\lambda_\kappa}{2}\mu_\kappa - \sinh\frac{\lambda_\kappa}{2}e^{2i\phi_\kappa}\mu_{-\kappa}{}^* \text{ for } \kappa \neq 0 \quad (16)$$

* See, for instance, COHEN M. H. and KEFFER F. *Phys. Rev.* **99**, no. 4, 1128–34 (1955).

VOL. 1 1956/57

where

$$\omega_\kappa \cosh\lambda_\kappa = \gamma[H - MN_z + H_e a^2\kappa^2 + 2\pi M \sin^2\theta_\kappa] \quad (17)$$

and

$$\omega_\kappa \sinh\lambda_\kappa = \gamma[2\pi M \sin^2\theta_\kappa] \quad (18)$$

The equations of motion are thereby brought into the form

$$\dot{\mu}_\kappa = i\omega_\kappa \mu_\kappa - \frac{1}{i\hbar}\alpha(\kappa)\sigma_0 + \frac{1}{i\hbar}\beta(\kappa)\sigma_0^* \quad (19)$$

$$\dot{\sigma}_0 = i\omega_0\sigma_0 - \frac{1}{i\hbar}\sum_\kappa \alpha^*(\kappa)\mu_\kappa + \frac{1}{i\hbar}\sum_\kappa \beta(\kappa)\mu_\kappa^* \quad (20)$$

where $$\omega_\kappa = \gamma\sqrt{(H - MN_z + H_e a^2\kappa^2)(H - MN_z + H_e a^2\kappa^2 + 4\pi M \sin^2\theta_\kappa)}$$

$$\omega_0 = \gamma[H - M(N_z - N_t)]$$

$$\alpha_\kappa = \cosh\frac{\lambda_\kappa}{2}\frac{S}{N}\sum_{jk} e^{-i\kappa\cdot r_j} b_{jk} + \sinh\frac{\lambda_\kappa}{2} e^{2i\phi_\kappa}\frac{2S}{N}\sum_{jk} e^{-i\kappa\cdot r_j} e_{jk} \quad (21)$$

$$\beta_\kappa = \sinh\frac{\lambda_\kappa}{2} e^{2i\phi_\kappa}\frac{S}{N}\sum_{jk} e^{-i\kappa\cdot r_j} b_{jk} + \cosh\frac{\lambda_\kappa}{2}\frac{2S}{N}\sum_{jk} e^{-i\kappa\cdot r_j} e_{jk}^* \quad (22)$$

The set of linear equations (19) and (20) enable the following question to be answered. Suppose at time $t = 0$ the spin system is disturbed so that the expectation value of the total transverse spin S_+ has an expectation value $\langle S_+(0)\rangle$; how will this expectation value develop in time? Using the method of Laplace transforms, and remembering that $S_+ = \sqrt{2SN}\,\sigma_0$, we obtain the result

$$\langle S_+(t)\rangle = \langle S_+(0)\rangle \frac{1}{2\pi i}\int \frac{e^{pt}\,dp}{p - i\omega_0 + \frac{1}{\hbar^2}\sum_\kappa\left(\frac{|\alpha(\kappa)|^2}{p - i\omega_\kappa} - \frac{|\beta(\kappa)|^2}{p + i\omega_\kappa}\right)} \quad (23)$$

where the integral is taken over a contour in the complex plane to the right of all poles of the integrand. The major contribution to the integral will come from the pole in the vicinity of $p = i\omega_0$.

We examine the sums in the denominator of (23), therefore, on the assumption that $p = p_1 + ip_2$, where $p_1 > 0$ and p_2 is close to ω_0. It is useful to allow the frequencies, ω_κ, to be complex and to write $\omega_\kappa = \Omega_\kappa + j\gamma_\kappa$. This assumption amounts to saying that the various spin waves for $\kappa > 0$ must physically be able to individually relax into the whole assemblage of spin waves or into the lattice vibrations. It will be remembered that the terms corresponding to such a mechanism have been deliberately dropped from eq. (19). Their presence is now simulated by assuming ω_κ to be complex. It is clear that the first of the sums is much greater than the second, because of the presence in the former of the denominator, $i(p_2 - \Omega)_\kappa + (p_1 - \gamma_\kappa)$, which becomes small at $p_2 = \Omega_\kappa \sim \omega_0$, provided that p_1 and γ_κ are small for $\Omega_\kappa \simeq \omega_0$. We assume this to be true in the sense that $\gamma_\kappa \ll \omega_0$. We, therefore, neglect the second sum. The first sum may be written as

$$\Sigma_1 - i\Sigma_2 = \sum_\kappa \frac{p_1 - \gamma_\kappa}{(p_2 - \Omega_\kappa)^2 + (p_1 - \gamma_\kappa)^2}|\alpha(\kappa)|^2 - i\sum_\kappa \frac{p_2 - \Omega_\kappa}{(p_2 - \Omega_\kappa)^2 + (p_1 - \gamma_\kappa)^2}|\alpha(\kappa)|^2$$

The real sum comes mainly from that part of κ-space in which $\Omega_\kappa \simeq p_2$; to first order its value does not depend upon $p_1 - \gamma_\kappa$ and is given by

$$\Sigma_1 = \frac{V}{4\pi}\int_0^\pi \left\{|\alpha(\kappa)|^2\kappa^2\left|\frac{\partial_\kappa}{\partial\omega_\kappa}\right|\right\}_{\omega_\kappa = \omega_0} \sin\theta_\kappa\, d\theta_\kappa \quad (24)$$

where V is the volume of the sample. The second term, Σ_2, on the other hand, depends markedly upon the behavior of γ_κ.

The pole of (23) now lies at $p = (-\Sigma_1/\hbar^2) + i\{\omega_0 + (\Sigma_2/\hbar^2)\}$ and lies to the left of the line $p = p_1 + ip_2$. Thus, (23) yields

$$\langle S_+(t)\rangle = \langle S_+(0)\rangle \exp\left\{i\left(\omega_0 + \frac{\Sigma_2}{\hbar^2}\right) - \frac{\Sigma_1}{\hbar^2}\right\} t = \langle S_+(0)\rangle \exp\left\{i(\omega_0 + \Delta\omega) - \frac{1}{\tau}\right\} t$$

where

$$\Delta\omega = \frac{1}{\hbar^2}\Sigma_2 \text{ and } 1/\tau = 1/\hbar^2\,\Sigma_1 \quad (25)$$

The disorder of the lattice has thus introduced a frequency shift $\Delta\omega$ into the motion of the total spin and furthermore causes the motion to relax

with a time constant τ. From (24), the relaxation time is independent of the value of γ_κ and it is clear that τ is to be associated with the process converting the uniform precession into degenerate spin waves. The expectation value of S_z does not change in such a process and we are dealing, therefore, with a Bloch type relaxation.

The frequency shift $\Delta\omega$ cannot be treated in this same fashion since it depends strongly on the behavior of γ_κ. If γ_κ were much smaller than ω_0 for all values of κ, the frequency shift would be rather large and hard to reconcile with experimental measurements of g-value and anisotropy providing equation (24) is simultaneously to predict a reasonable relaxation time. On the other hand, it is not unreasonable to suppose that γ_κ increases very fast for large κ both by interactions with the spin system and with the lattice. We shall not at this point attempt to calculate $\Delta\omega$, but it is possible that the disorder of the lattice may make a substantial contribution to the g-value and anisotropy.

In order to proceed with the calculation of $1/\tau$ on the basis of eq. (24) we must evaluate $|\alpha(\kappa)|^2$. If it is assumed that the fluctuation in the dipolar coefficient between all pairs of ions is completely uncorrelated, one obtains

$$|\alpha(\kappa)|^2 = \frac{1}{N}\frac{9}{4}(g\beta H_p)^2\frac{1}{Z}\sum_{n.n.}\left|\cosh\frac{\lambda\kappa}{2}(3\gamma_{jk}^2-1)+\right.$$
$$\left.+\sinh\frac{\lambda\kappa}{2}e^{2i\phi_\kappa}(\alpha_{jk}-i\beta_{jk})^2\right|^2 \qquad (26)$$

where the sum is taken over nearest neighbors, N is the total number of spins and H_p^2 is related to the mean square fluctuation of D_{jk} by

$$H_p^2 = Z\left(\frac{S}{g\beta}\right)^2\langle\Delta D_{jk}^2\rangle \qquad (27)$$

Because the model we have adopted is greatly simplified compared to the magnetic structure of an inverted spinel, it seems unrealistic to adopt any particular crystal symmetry in order to evaluate the sum in eq. (26). Instead we shall replace the sum by an isotropic average and obtain

$$|\alpha(\kappa)|^2 = \frac{1}{N}\frac{9}{5}(g\beta H_p)^2\left(\cosh^2\frac{\lambda\kappa}{2}+\frac{2}{3}\sinh^2\frac{\lambda\kappa}{2}\right) \qquad (28)$$

In so doing we naturally forego any possibility of discussing a possible anisotropy of line width.

Equation (28) may now be introduced into eq. (24). With suitable changes in the variable of integration one obtains

$$\Delta H = \left(\frac{V}{Na^3}\frac{3}{20\pi}\right)H_p^2\frac{\sqrt{4\pi M}}{H_e^{3/2}}I\left(n_t, \frac{H}{4\pi M}\right) \qquad (29)$$

where the result is expressed by $\Delta H = 2/\gamma\tau$, as more related to the experimental observations, and the integral I is given by

$$I\left(n_t, \frac{H}{4\pi M}\right) = \frac{H}{4\pi M}\int_1^{\text{Limit}}\left[\frac{1}{x}+\frac{5}{2}\left(1+\frac{1}{x^2}\right)\right]\times$$
$$\times\sqrt{\frac{1-\left(1-n_t\frac{4\pi M}{H}\right)x}{1+\frac{4\pi M}{H}x-x^2}}\,dx \qquad (30)$$

The upper limit of the integral is fixed by the vanishing of either the numerator or denominator of the square root in the integrand. The quantity n_t is $N_t/4\pi$. The integral has been evaluated numerically as a function of n_t and $H/4\pi M$, and is expressed graphically in Fig. 2.

3. CONCLUSION

In this final section we shall consider the results of the theory as embodied in eq. (29) and Fig. 2. We first examine the magnitude of the line width predicted for the disorder scattering process. The expression for ΔH is a product of several factors of which the first is numerical and of order of 1/25. The quantity H_p must have roughly the magnitude of the pseudo-dipolar fields and be of the order of 10^5 oersteds. A typical value for $\sqrt{(4\pi M)}$ is about 70, and we may suppose the exchange field H_e to be about 10^6 oersteds. For a spherical sample, a representative value of $I(n_t, H/4\pi M)$ is 1·5. From these numbers we find ΔH approximately equal to 40 oersteds. This value is close enough to observed magnitudes so that it is highly probable that disorder scattering plays a significant role in the experimentally determined

VOL. 1 1956/5

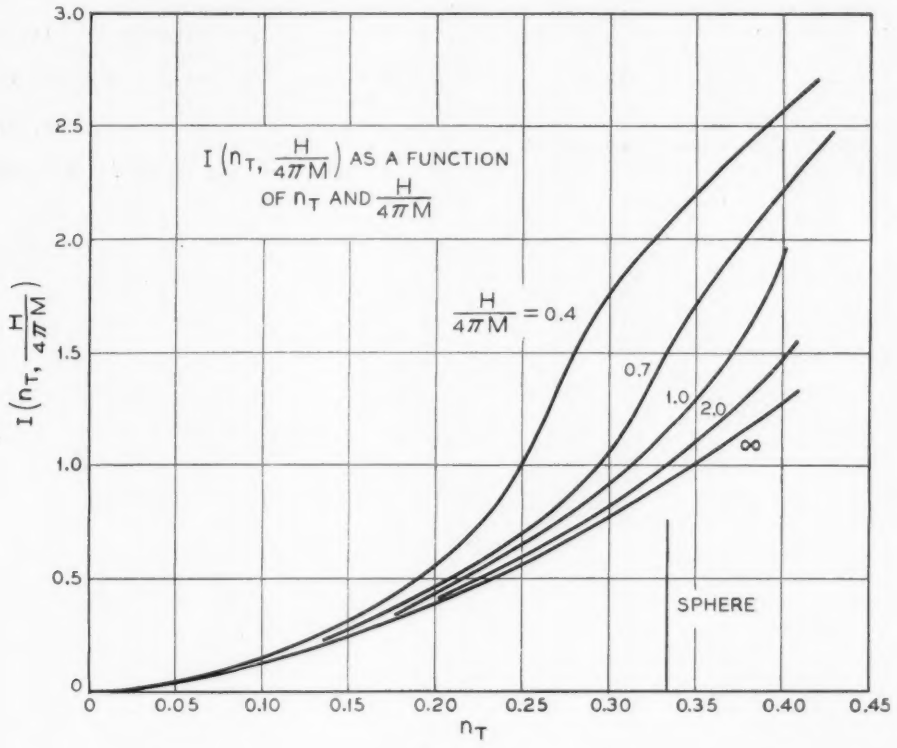


FIG. 2.

line widths. If this is so, the measured line widths should in corresponding degree show the following behavior,

(1) ΔH should increase roughly as $\sqrt{M}$. This effect would be best observed by measuring a given sample at different temperatures.

(2) ΔH should decrease markedly with the transverse demagnetization factor. Thus, the line width should be much less as measured in a thin disk normally magnetized than in a sphere of the same material.

(3) All else being equal, materials of high Curie temperature should have narrower line widths.

(4) Any ordered magnetic material should have a narrower line width. It would, therefore, be very interesting to obtain line width measurements on magnetite above and below the order-disorder transformation.

It is interesting also to consider the line shape to be expected from the disorder scattering mechanism. Near resonance the line should be closely Lorentzian corresponding to the relaxation time τ. On the high frequency side of the line the absorption should be greater than expected because of the increasing density of states in the spin wave spectrum at values of κ outside the degenerate manifold. On the other hand, the low frequency side of the line should show a correspondingly smaller absorption decreasing to zero for $\omega \leqslant \gamma(H-N_zM)$.

REFERENCES

1. DILLON J. F., GESCHWIND S. and JACCARINO V. *Phys. Rev.* **100**, no. 2, 750–2 (1955).
2. TANNENWALD P. E. *Phys. Rev.* **100**, no. 6, 1713–4 (1955).
3. GALT J. K. *Phys. Rev.* **80**, no. 4, 744–8 (1950); *Phys. Rev.* **99**, no. 4, 1203–10 (1955).

4. AMENT W. S. and RADO G. T. *Phys. Rev.* **97,** no. 6, 1558–66 (1955).
5. KITTEL C. and ABRAHAMS E. *Rev. of Mod. Phys.* **25,** no. 1, 233–8 (January 1953).
6. VANVLECK J. H. *Phys. Rev.* **78,** no. 3, 266–74 (1950).
7. KEFFER F. *Phys. Rev.* **88,** no. 3, 686–7 (1952).
8. BLOEMBERGEN N. and WANG S. *Phys. Rev.* **93,** no. 1, 72–83 (1954).
9. KASUYA T. *Progress of Theoretical Physics* **12,** no. 6, 802–3 (December 1954).
10. HOLSTEIN T. and PRIMAKOFF H. *Phys. Rev.* **58,** no. 1098–1113 (1940).
11. HERRING C. and KITTEL C. *Phys. Rev.* **81,** no. 5, 869–80 (1951).
12. ANDERSON P. W. and SUHL H. *Phys. Rev.* **100,** no. 6, 1788–9 (1955).
13. KITTEL C. *Phys. Rev.* **73,** no. 2 155–61 (1948).

VOL.
1
1956/5

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 137–142.

HYDROGEN ATOM IN A STRONG MAGNETIC FIELD

Y. YAFET, R. W. KEYES AND E. N. ADAMS

Westinghouse Research Laboratories, Pittsburgh 35, Pennsylvania

(*Received* 24 *July* 1956)

Abstract—The energy levels and wave functions are studied for a hydrogen atom in a very strong magnetic field. The contraction of the electronic-charge distribution in a plane perpendicular to the magnetic field results in an increase in ionization energy at high fields. Conditions are favorable for observing the effect in the case of impurity states in a semiconductor of high dielectric constant and low effective mass, to which the theory is applied in detail. Changes in the optical and electrical properties are predicted.

1. INTRODUCTION

SHALLOW electronic energy states associated with impurity levels in semiconductors can sometimes be understood by means of a hydrogen model of the impurity.[1, 2] In such cases one takes as the wave functions and energy levels for the various electronic impurity states the corresponding wave functions for a hydrogen atom in which the electron has a mass m^*, the effective mass of the electron in the crystal, and the nucleus has a charge e/K, the normal protonic charge decreased by the high-frequency dielectric constant of the medium.

The ratio of the Zeeman splitting energy to the Rydberg in such hydrogen-like states is strongly dependent on the values of m^* and K. In particular for materials in which the electron effective mass is small and the dielectric constant large, this ratio becomes very much larger than usual, and may, in fact, greatly exceed unity under practically attainable experimental conditions.

We have made a study of the theory of the hydrogen atom in such very strong magnetic fields. We will present our results in terms of the dimensionless parameter

$$\gamma = \hbar\omega/2Ry$$

where ω is the cyclotron frequency of a free carrier in the magnetic field and Ry is the Rydberg constant. γ is the ratio of $(\hbar\omega/2)$, the zero point energy of a free carrier in the magnetic field, to Ry, the energy of a carrier in the lowest state of the hydrogen atom placed in a field-free region.

For $\gamma \ll 1$ the effect of the magnetic field on the hydrogen levels is well known. However, if γ is increased to values of the order of unity or more the treatment of the magnetic terms as a merely perturbative influence on the hydrogen atom becomes invalid. For $\gamma \sim 1$ it would presumably be very difficult to find either wave functions or energy levels by means of any procedure based on the use of free atom wave functions.

For fields so great that $\gamma \gg 1$, however, the situation becomes simple again. In this case it is feasible to find the wave functions and energy levels by an expansion in terms of the free carrier wave functions using a procedure conceived in the spirit of perturbation theory.

Before making any quantitative or mathematical remarks we will describe in purely qualitative terms just what we expect to happen to the hydrogen atom as the magnetic field is increased so that γ takes on values of unity or greater. Since the magnetic forces are centripetal in a plane perpendicular to the field H, they tend to compress the atom in the transverse dimension. For weak fields ($\gamma \ll 1$) the magnetic forces are not comparable to the coulomb forces so the effect is chiefly to lower the energy of a state with a favorably aligned magnetic moment compared to that of one with an unfavorably aligned moment, without, however, having much effect on the wave function describing the state. On the other hand, as γ approaches unity the magnetic forces have a larger and larger effect on the electronic wave function. The increase of the equatorial forces engenders a corresponding decrease of the equatorial diameter, so the atom

becomes increasingly ovoid in shape. The calculations to be given later indicate that for $\gamma \sim 1$ the ground state of the atom is egg-shaped, for $\gamma \sim 100$ cigar-shaped.

As the atom elongates it shrinks in all dimensions. Thus there is a steady increase in the magnitude of the (negative) coulomb energy. As a consequence the ionization energy becomes greater as the field is increased. This increase of binding energy, which provides a basis of one experimental means of observing the effects, will be calculated in a later section.

In the following sections we will present the detailed theory of these effects. In the first section we describe means of obtaining the wave functions and energy levels in the limits $\gamma \gg 1$, and give a descriptive account of these. In the second section we present a variational calculation of the ground state energy and describe its dependence on γ. In the third section we discuss the thermal equilibrium of an ensemble of hydrogen atoms at elevated temperatures and for the condition $\gamma \gg 1$. In the fourth section we discuss the experimental situation that obtains in various known semiconductors with a view to determining which cases are especially favorable for studying the effects experimentally.

1. *Energy Levels and Wave Functions*

The Hamiltonian for the motion of a hydrogen atom in a homogeneous uniform magnetic field is

$$\mathcal{H} = \frac{p^2}{2m} + \frac{1}{2}\hbar\omega L_z + \frac{1}{8}m\omega^2(x^2+y^2) - \frac{Ze^2}{r} \tag{1.1}$$

Here L_z is the dimensionless operator for the component of orbital angular momentum along the magnetic field. Introducing the Rydberg ($mZ^2e^4/2\hbar^2$) as unit of energy, the Bohr radius $\hbar^2/mZe^2$ as unit of length, and using the parameter γ described above, we can re-express (1.1) as

$$\frac{\mathcal{H}}{Ry} = -\nabla^2 + \gamma L_z + \frac{\gamma^2}{4}(x^2+y^2) - \frac{2}{r} \tag{1.2}$$

Here x, y, and r are dimensionless and ∇ is the dimensionless gradient operator in these variables.

We begin by discussing briefly the eigenfunctions and eigenvalues of the operator obtained by dropping the coulomb term from the right-hand side of (1.2). Because of a fundamental degeneracy in the problem these eigenfunctions may be classified in various ways.[3]† We will employ the classification in which one quantum number is M, the eigenvalue of L_z. In that case the zero order eigenfunctions are

$$\psi_{MNk}^{\circ} = \phi_{NM}(x,y)e^{ikz} \tag{1.3}$$

and the eigenvalues

$$E_{NMk} = 2(N+\tfrac{1}{2})\gamma + E_z, \quad N \geqslant M. \tag{1.4}$$

In the above equations N is the principal quantum number, k is the wave number of motion along the magnetic field, E_z is the kinetic energy of motion along the field, E_z is $\hbar^2k^2/2m$ in dimensional units. The complete energy-level diagram for $k = 0$ is shown in Fig. 1. It is noteworthy that M does not occur directly in the energy.

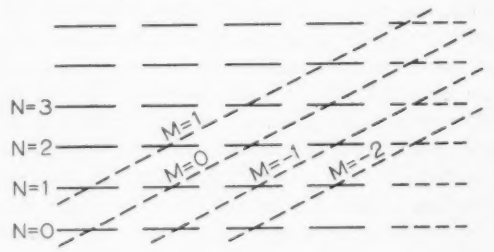


FIG. 1. Energy level diagram for free carriers in magnetic field.

The states φ_{NM} are states for which the electronic charge distribution is limited to a cylindrical region of radius $(2N-2M+1)\hbar/m\omega^{\frac{1}{2}}$ about the origin. Thus these states are localized in the xy plane in such a way as to be useful in constructing atomic-type wave functions. We will describe now a procedure for constructing from the functions φ_{NM} functions which we will designate as $\psi_{NM\alpha}(xyz)$, these $\psi_{NM\alpha}$ being such as to permit a perturbative solution of the strong field coulomb problem. With this goal in mind we define the one-dimensional potential

$$V_{NM}(z) = -\int dx\,dy\,|\phi_{NM}|^2\frac{e^2Z}{r}. \tag{1.5}$$

We will first show that in the limit $\gamma \to \infty$ the exact eigenfunctions of (1.2) have the form

$$\psi_{NM\lambda} = \phi_{NM}(x,y)f_{NM\lambda}(z). \tag{1.6}$$

† While other authors have studied these functions, the reference cited is particularly useful.

VOL. 1 1956/57

The functions $f_{NM\lambda}(z)$ are the solutions of the one-dimensional Schrödinger problem with the potential $V_{NM}(z)$. They satisfy*

$$\left(-\frac{d^2}{dz^2}+V_{NM}(Z)-\epsilon_{NM\lambda}\right)f_{NM\lambda}(z)=0. \tag{1.7}$$

The exact energy levels become

$$E_{NM\lambda}=(2N+1)\gamma+\epsilon_{NM\lambda}. \tag{1.8}$$

These assertions are easily seen to be correct in virtue of the following two facts.

(*a*) The Hamiltonian has off diagonal matrix elements only between states of the same M. However, the construction of the ψ_{NM} is such that there can be no matrix elements between states of the same NM and different λ.

(*b*) From (*a*) and a study of Fig. 1 it is easily seen that the only states which are mixed have different N, hence energies differing by approximately an integral multiply of $\hbar\omega$. The admixture of states in such a case is small like $1/\gamma$, and vanishes as $\gamma\to\infty$.

We now can characterize the various energy states of the hydrogen atom in a very strong magnetic field. We will be interested here in discussing only those low-lying states with $N=0$, the states $\varphi_{0M}f_{0M\lambda}$ For these states LIPPMAN and JOHNSON[3] have shown that the charge distribution in the xy plane is confined to a washer-shaped region. The mean radius of the washer is $\sim[(2|M|+1)\hbar/m\omega]^{\frac{1}{2}}$ and the width of the washer is $\sim(\hbar/m\omega)^{\frac{1}{2}}$.

In the states for which $M\neq 0$ the electron is kept away from the origin by the action of the centrifugal potential, which is proportional to M^2. That is why the mean electronic distance from the origin is proportional to $(2|M|+1)^{\frac{1}{2}}$ in the various low-lying states. The state of $M=0$ has the smallest mean radius in the xy plane. Therefore, it must be the ground state in the strong field limit, since the coulomb energy will have its greatest (negative) value for the smallest state. By the same token the states for very large M will have a negligible coulomb energy, and will thus be almost indistinguishable from free carrier states in the magnetic field.

The ionization energy in the strong field case will be the energy difference between the energy of the lowest $N=0$, $M=0$ state, and that of the lowest $N=0$, $M=-\infty$ state.

2. *Ground State and Ionization Energy*

In this section we will make a variational estimate of the ground-state binding energy as a function of magnetic field strength. The ground state is an $M=0$ state, hence is isotropic about the magnetic field. A normalized variational trial function of the right symmetry is

$$\psi=(2^{3/2}a_\perp^2a_\parallel\pi^{3/2})^{-\frac{1}{2}}\exp\left(-\frac{x^2+y^2}{4a_\perp^2}-\frac{z^2}{4a_\parallel^2}\right). \tag{2.1}$$

A straightforward computation gives for the trial value of the energy

$$E=\frac{1}{2a_\perp^2}\left(1+\frac{\epsilon^2}{2}\right)+\frac{\gamma^2a_\perp^2}{2}-\\-\frac{\epsilon\sqrt{2}}{a_\perp\sqrt{\pi}\sqrt{(1-\epsilon^2)}}\log\frac{1+\sqrt{(1-\epsilon^2)}}{1-\sqrt{(1-\epsilon^2)}} \tag{2.2}$$

in which we have introduced $\epsilon=a_\perp/a_\parallel$. Minimizing the expectation value of the energy with respect to ϵ and $a_\perp$ we obtain the two equations:

$$-\frac{1}{a_\perp^3}\left(1+\frac{\epsilon^2}{2}\right)+\gamma^2a_\perp+\\+\frac{\epsilon\sqrt{2}}{a_\perp^2\sqrt{\pi}\sqrt{(1-\epsilon^2)}}\log\frac{1+\sqrt{(1-\epsilon^2)}}{1-\sqrt{(1-\epsilon^2)}}=0 \tag{2.3}$$

$$\frac{\epsilon}{2a_\perp^2}+\frac{1}{a_\perp}\cdot\sqrt{\frac{2}{\pi}}\left[\frac{2}{1-\epsilon^2}-\right.\\\left.-\frac{1}{(1-\epsilon^2)^{3/2}}\log\frac{1+\sqrt{(1-\epsilon^2)}}{1-\sqrt{(1-\epsilon^2)}}\right]=0. \tag{2.4}$$

It does not seem feasible to solve (2.3) and (2.4) explicitly for $a_\perp$ and ϵ in terms of γ. However, it is

* It has been pointed out to us that Schiff and Snyder[a] have considered the problem of the effect of a magnetic field on excited states in a Coulomb potential, and have shown that even in a weak field there are highly excited states with energy given by (1.7) and (1.8). These authors were, however, concerned with fields orders of magnitude smaller than those which we are considering here.

[a] L. I. SCHIFF and H. SNYDER, *Phys. Rev.* **55**, 59 (1939).

straightforward to obtain numerical solutions for given values of γ. Fig. 2 shows a graph of the results of such calculations. In it are plotted the lengths giving the spatial extent of the wave function; together with $a_0 = (\hbar c/eH)^{\frac{1}{2}}$, the spatial extent of the wave function of a free carrier in the magnetic field. Note that a_0 approaches $a_\perp$ at large γ.

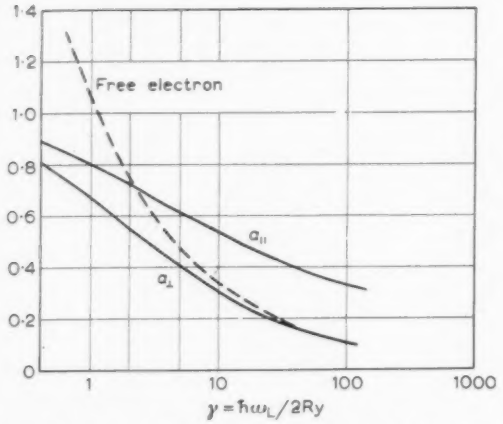


FIG. 2. The best values for the parameters of the wave function (2.1), together with a_0, the equivalent of a for a free electron, as functions of γ. The unit of length is the best value of $a_\perp$ and $a_\parallel$ in zero magnetic field, equal to 0.94 of the appropriate Bohr radius.

The ionization energy is the difference between the energies of the lowest bound state and the lowest free state. Our variational lower bound for it is given by

$$E_B = \gamma - \frac{1}{2a_\perp{}^2}\left(1+\frac{\epsilon^2}{2}\right) - \frac{\gamma^2 a_\perp{}^2}{2} + \\ + \frac{\epsilon\sqrt{2}}{a_\perp\sqrt{\pi}\sqrt{(1-\epsilon^2)}} \log \frac{1+\sqrt{(1-\epsilon^2)}}{1-\sqrt{(1-\epsilon^2)}} \quad (2.5)$$

Fig. 3 shows a plot of E_B vs. γ as obtained from (2.5) and the numerical values of $a_\perp$ and $a_\parallel$ obtained above.

No estimate of the error of the variational solution was made at the higher fields, although it would be possible to make an exact solution there if necessary. The error in the ionization energy at zero field is 15 per cent. We think it likely that the percentage error is no worse at high fields, since

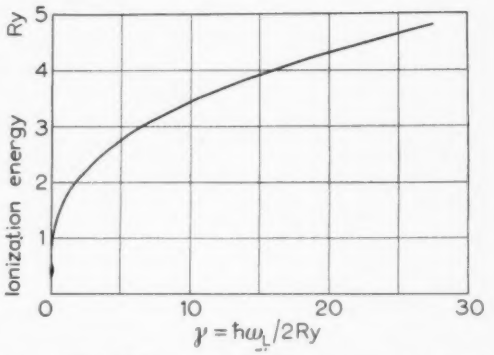


FIG. 3. The difference in energy between the lowest state of the conduction band and the lowest state bound to the impurity atom, as a function of the magnetic field strength. The energy of the bound state is calculated by the variational method described in the text.

the variational form is more nearly correct at the high fields.

We have attempted to get an analytic approximation to the binding energy at high fields. An approximation which gives errors of the order of 10 per cent or less when $\gamma > 15$ is

$$E_B = \frac{1}{4}\epsilon^2\gamma + \sqrt{\frac{8}{\pi}}\epsilon\gamma^{\frac{1}{2}} \quad (2.6)$$

with ϵ given by the solution of

$$\frac{\sqrt{\pi}\gamma^{\frac{1}{2}}}{2^{3/2}} = \frac{2}{\epsilon}\left(\ln\left[\frac{2}{\epsilon}\right] - 1\right). \quad (2.7)$$

3. *Feasibility of Experimental Test*

The magnetic field for which $\gamma = 1$ is given by the relation

$$H = 2\times 10^9 \left(\frac{m^* Z}{m_0}\right)^2 \text{ oers.} \quad (3.1)$$

in which m^* is the electron effective mass, m_0 is the free electron mass, and Z the nuclear charge. For a natural hydrogen atom (3.1) yields a magnetic field value of 2×10^9 oers, a value far in excess of any that is experimentally feasible. Thus the experimental study of the strong field region will of necessity be limited to materials in which the quantity (Zm^*/m_0) is much smaller than unity.

For hydrogenic impurity states in crystals the

VOL. 1 1956/5

effective nuclear charge Z is the reciprocal of the high frequency dielectric constant K.

The electron effective mass depends on the nature of the crystal. We will denote it as $m^* = m_0/\alpha$. Thus when α is large the effective electron mass is small.

We can now rewrite eq. (3.1) for the critical strong field condition as

$$H = 2\times 10^9 \text{ oers } (\alpha K)^{-2} \qquad (3.2)$$

Certain semiconducting materials naturally suggest themselves as especially suitable for study in a strong magnetic field both because they have favorable values of αK and because they can be readily prepared in high purity crystalline condition. We will consider two examples, In Sb and Ge, although there are probably others which would be of experimental interest.

For InSb[4, 5] the value of K is 16 and the value of α approximately 77. Thus the product αK is about 1200 and the critical field is about 1300 oersted.

For Ge the value of K is also 16 and the value of α approximately 10 depending somewhat on orientation. Thus the αK is about 160 and the critical field is about 77,000 oers.

The highest magnetic fields currently available are approximately 10^5 oersted steady and up to 10^6 oersted in short pulses. Clearly for experiments on InSb either the steady or pulsed field method is adequate to attain strong field limit $\gamma \gg 1$. For Ge, on the other hand, the static fields are not adequate to attain the strong field limit, and are barely sufficient to attain the condition $\gamma = 1$. Thus pulsed fields would be required to attain the condition $\gamma \gg 1$.

Two methods of observing the effect of the strong magnetic field immediately suggest themselves, viz. observations of changes in optical absorption and electrical conductivity.

The effect of the strong magnetic field on the optical absorption in Ge might be feasible to observe. The Lyman series associated with impurity centers in Si has been observed by BURSTEIN, and the similar series in Ge is within the range of available experimental technique. This series could presumably be studied as a function of magnetic field. A magnetic field for which $\gamma = 1$ would effect quite a large change in the spectrum.

The optical absorption method is more difficult to use in InSb. In the first place the absorption line will be in that part of the submillimeter region where technical problems are most difficult. In the second place one would have to work with concentrations of impurities which are very low, so there will be a signal to noise ratio problem.

Under certain circumstances it is possible to use an electrical method of detecting the strong field change of binding energy. The experiment consists in observing the magnetoresistance and/or Hall effect at an appropriately low temperature and as a function of magnetic field. If the temperature is properly chosen an increase in the binding energy can cause a large decrease in the population of carriers thermally free of traps. Thus as the field is changed at constant temperature there should be a large change in the conductivity because of the change in carrier concentration.

The electrical method is capable of giving only limited information concerning the effect of the field on the impurity state. However, it has the advantage of being very sensitive, so that very low carrier concentrations are quite adequate for experimental purposes.

In order to consider what information about the binding energy could be obtained from electrical measurements of the sort mentioned above, we have made a study of the statistical distribution of carriers among the various states in the strong field limit. A correct calculation of these effects is actually rather complicated because of the peculiar spectrum of states in the coulomb field. Thus in this paper we will content ourselves with a description of the qualitative factors which are involved in the carrier deionization.

Roughly speaking the degree of ionization of carriers depends on the number of bound states, the binding energy of the bound states, and the density of unbound states. Both of the last two quantities increase in a strong field. Since these two changes are in such a direction as to cause opposite effects on the number of ionized carriers, the sign of net change will depend on which factor is dominant.

We have previously discussed the behavior of the binding energy with magnetic field, and seen that it increases very strongly with increasing H in the neighborhood of $\gamma \sim 1$, but more weakly for $\gamma \gg 1$. The number N_T of unbound states which is "thermally available" is approximately given by

$$N_T = \frac{eH}{\pi^2\hbar^2 c}(2mkT)^{\frac{1}{2}}. \tag{3.3}$$

Thus the density of states in the high field region is directly proportional to the magnetic field.

In the approximation of nondegenerate statistics the number of carriers which are in unbound states is given roughly by

$$n_{\text{free}} = \frac{nN_T}{N_T + N_I \exp(E_B/kT)} \tag{3.4}$$

In which N_I is the number of impurity centers, E_B is the binding energy, and n is the total number of carriers. Whether n_{free} decreases or increases with increasing H depends on whether N_T or $N_I \exp(E_B/kT)$ increases more strongly.

An examination of eq. (3.4) in the light of the calculated behavior of E_B as given in section II and N_T as given in eq. (3.3) shows that either ionization of deionization may occur. If the temperature is sufficiently low the increase in the term $N_I e^{E_B/kT}$ will always exceed the increase in N_T so deionization will occur. On the other hand for sufficiently high temperature it is clear that the increase in N_T will be greater. An intermediate case is also possible in which the effect is ionization at weak fields followed by deionization at stronger fields. Such a case is shown in Fig. 4, which was computed on the basis of a somewhat more accurate analysis than we are giving here.

An attempt to observe some of these effects has been made by KEYES and SLADEK. Their preliminary results indicate that effects of the sort predicted here do occur. However, there are certain features of their experiment which do not conform to the model assumed in our calculations, so we are unable to directly apply our theory to the interpretation of their data.

Germanium seems to be most favorable for doing an experiment which conforms to the model used in our theory. The possibility of seeing an effect using fields no greater than 100,000 gauss and electrical methods is not excluded, although the experiment is probably marginal. It should be quite feasible to see very big effects on electrical resistance by means of experiments using pulsed fields provided that the relaxation processes are fast enough for carriers to be trapped during the period of duration of the pulse.

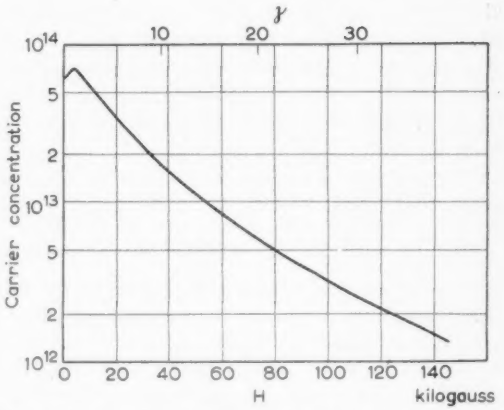


FIG. 4. An example of the way in which the concentration of carriers in the conduction band varies with magnetic field according to our model.

$N_D = 8{\cdot}25 \times 10^{14}$. $m^*/m = 1/50$.

$N_A = 0$. $1Ry^* = 12{\cdot}3°\text{K}$.

$K = 16$. $T = 4{\cdot}2°\text{K}$.

REFERENCES

1. PEARSON G. L. and BARDEEN J. *Phys. Rev.* **75,** 865 (1949).
2. KOHN W. and LUTTINGER J. M. *Phys. Rev.* **98,** 915 (1955).
3. JOHNSON M. H. and LIPPMANN B. A. *Phys. Rev.* **76,** 828 (1949).
4. DRESSELHAUS G., KIP A. Z., KITTEL C., and WAGONER G. *Phys. Rev.* **98,** 556 (1955).
5. BRIGGS H. B., CUMMINGS R. Z., HROSTOWSKI H. J., and TANENBAUM M. *Phys. Rev.* **93,** 912 (1954).

VOL. 1 1956/5

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 143–145.

EFFECT OF A MAGNETIC FIELD ON DONOR IMPURITY LEVELS IN InSb

ROBERT W. KEYES AND R. J. SLADEK

Westinghouse Research Laboratories, Pittsburgh 35, Pennsylvania

(*Received* 24 *July* 1956)

Abstract—In an attempt to observe the effect described in the preceding paper, we have measured the Hall constant of two specimens of InSb in very strong magnetic fields. The anticipated effect is found to take place, although the details are modified by interaction between states on different impurity atoms. In addition, it is found that the effect of the magnetic field on the electron mobility is larger than is expected on the basis of the usual theories of magnetoresistance.

INTRODUCTION

In the preceding paper[1] a theory of the effect of a strong magnetic field on the energy levels and wave functions of a hydrogen-like impurity state in a semiconductor is presented. It is concluded that the change in the energy levels of such an impurity state should be observable in *n*-type InSb as a decrease in the number of electrons in the conduction band in high magnetic fields at low temperatures. In an attempt to observe this effect we have measured the conductivity and the Hall constant of two specimens of InSb as a function of magnetic field strength up to fields of 80 kilogauss at 4·2°K.

The simple hydrogenic model of the donor states presented does not apply to InSb of available purity. Even at temperatures far below the Rydberg (9°K) the electrons do not fall from the conduction band into bound states.[2] This is apparently a result of overlap of electronic wave functions centered on different impurity atoms.[3] In InSb with 2×10^{14} excess donors per cm^3 the average distance between excess donor atoms is only three times the reduced Bohr radius. Nevertheless, the possibility of observing the freeze-out of carriers in a strong magnetic field still exists, because the reduction in the spatial extent of the wave functions by the magnetic field may be expected to decrease the overlap between adjacent functions.

EXPERIMENT

Samples of InSb were provided for us by Dr. S. W. Kurnick of the Chicago Midway Laboratories. Two specimens were used for the experiment. One had an excess donor concentration of 2×10^{16} cm^{-3} and the other a concentration of 4×10^{14} cm^{-3}. The samples were cut to dimensions $0{\cdot}1\times0{\cdot}2\times1{\cdot}0$ cm^3, the surfaces were etched, and potential and current leads were attached with In–Sn solder.

The high magnetic field facilities of the Cryogenics Branch of the Naval Research Laboratory were made available to us through the courtesy of Dr. R. T. Webber. Using these facilities, we have measured the Hall constant and the conductivity of the specimens as a function of magnetic field up to a field of 82 kilogauss at 4·2°K. The results of this measurement are presented in Figs. 1 and 2 as plots of number of carriers and of mobility as functions of magnetic field.

DISCUSSION OF NUMBER CHANGE

Fig. 1 shows that in the 4×10^{14} sample the number of carriers begins to decrease sharply when the field is raised above 20 kilogauss, and at the highest field is fifty times smaller than at zero field. This change is far greater than those obtained from the ordinary mechanisms which lead to a field dependent Hall constant[4,5] and we believe that it results from the creation of bound impurity states in the way described above.

Fig. 1 also shows that the large decrease in carrier concentration does not occur in the 2×10^{16} sample. This is consistent with our interpretation, since in the 2×10^{16} sample fifty times as many electronic wave functions must be placed in a given volume as in the 4×10^{14} sample, and the overlap of the wave functions will persist to much higher

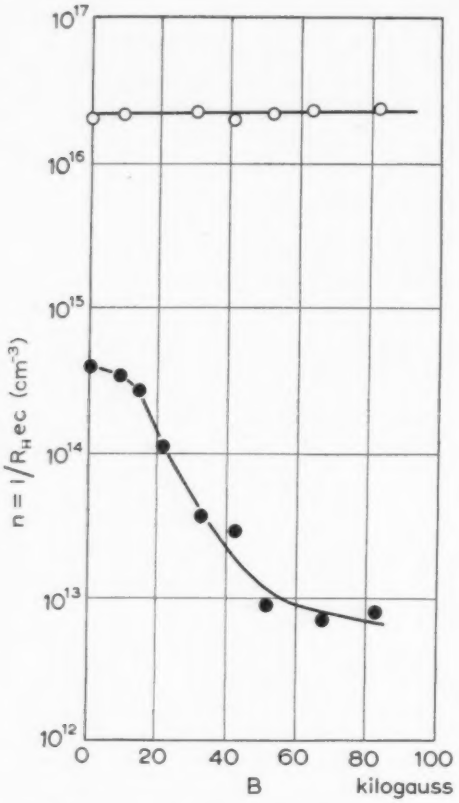


FIG. 1. The number of electrons, calculated from the Hall constant, as a function of magnetic field for two samples of InSb.
n-type InSb.
○ ~ 10^{16} sample.
● ~ 10^{14} sample.
4·2°K.

magnetic fields. On the other hand, any field dependence of the Hall constant which had its origin in a property of the conduction band states would probably not be greatly different in the two specimens.

We are not in a position to make a detailed comparison of our results with the calculated binding energy,[1] because the calculation does not take into account the effects of overlap of wave functions on different donors. These effects we have seen to be of great importance. Therefore the simple theory cannot be expected to be quantitatively applicable to the results of our experiment.

DISCUSSION OF THE MOBILITY

Fig. 2 shows that in our samples there is a quite large decrease in mobility at high magnetic fields The decrease is much larger than the usual classical theories of mobility[6] predict. Further, these theories always predict a saturation of the magnetoresistance effect[4,6] at high values of the dimensionless parameter μB. In our experiment on the 2×10^{16} sample, values of μB as high as 60 were reached, but the mobility does not appear to reach a saturation value. Thus we conclude that the usual theories of magnetoresistance are not adequate to account for our data.

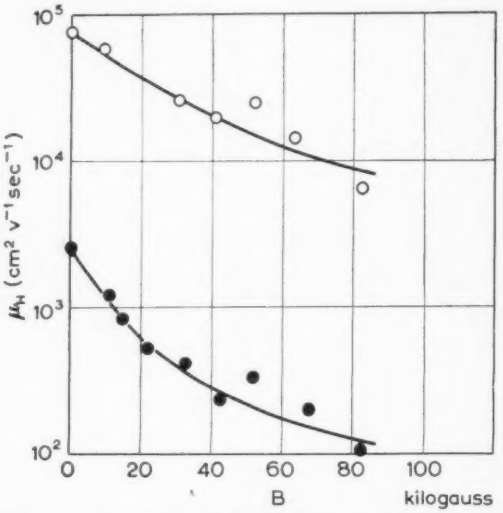


FIG. 2. The mobility of the electrons as a function of magnetic field in samples of InSb with excess electron concentrations of 2×10^{16} cm^{-3} and 4×10^{14} cm^{-3}.
n-type InSb.
○ ~ 10^{16} sample.
● ~ 10^{14} sample.
4·2°K.

The large difference in the zero field mobility of the two samples deserves comment. Both samples are degenerate at 4·2°K, so that the electrons which participate in the conduction process have kinetic energies corresponding to the Fermi energy of the sample in question. If the scattering is by ionized impurities, the relaxation time varies approximately as the 3/2 power of the electronic energy. For comparable concentration of scattering centers, the sample with the higher degeneracy temperature therefore has a higher mobility. Since the

Fermi energy varies as the 2/3 power of the electron concentration, the mobilities are proportional to the concentration. For our samples which have a concentration ratio of 50, the ratio of the mobilities is 30. This suggests that the number of scattering centers is about the same in our two samples, and that the low electron concentration in the 4×10^{14} sample is achieved by compensation of donor and acceptor impurities.

CONCLUSIONS

1. We find that the energy difference between a donor impurity level and the bottom of the conduction band can be increased by the application of a strong magnetic field, and the distribution of electrons between the conduction band and impurity states can thereby be altered. 2. The effect of a strong magnetic field on the mobility is greater than can be accounted for by the usual theories of magnetoresistance.

Acknowledgements—We are grateful to Dr. T. T. WEBBER and the Cyrogenics Branch of the Naval Research Laboratory for making available to us the facilities for the experimental work described in this paper. We wish to thank the Chicago Midway Laboratories and Dr. S. W. KURNICK for supplying the InSb used in these experiments. We are indebted to Dr. E. N. ADAMS, who suggested to us the existence of the effects described herein, and to Dr. Y. YAFET, for discussions of the interpretation. Since this experiment was performed, Dr. H. P. R. FREDERIKSE has observed a similar effect, and we thank Dr. FREDERIKSE for discussing his results with us.

REFERENCES

1. YAFET Y., KEYES, R. W., and ADAMS E. N. Preceding paper.
2. HROSTOWSKI H. J., MORIN F. J., GEBALLE T. H., and WHEATLEY G. H. *Phys. Rev.* **100,** 1672 (1955).
3. STERN F. and TALLEY R. M. *Phys. Rev.* **100,** 1638 (1955).
4. SWANSON J. A. *Phys. Rev.* **99,** 1799 (1955).
5. WILLARDSON R. K., HARMAN L., and BEER A. C. *Phys. Rev.*. **96,** 1512 (1954).
6. WILSON A. H. *Theory of Metals*, Cambridge University Press, Cambridge (1953).

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 146–158.

COMPRESSION OF SOLIDIFIED GASES TO 20,000 kg/cm² AT LOW TEMPERATURE*

JOHN W. STEWART

University of Virginia, Charlottesville, Virginia

(*Received* 28 *June* 1956)

Abstract—The relative volume decrease as a function of pressure has been measured for solid He, H_2, D_2, and Ne at 4°K, for N_2 and A at 65°K, and for A at 77°K by the direct piston displacement method of BRIDGMAN. The results for He and N_2 are new, while for the other gases earlier work has been confirmed and extended to higher pressure. Solid helium is found to be the most compressible solid yet measured, being reduced in volume by a factor of 2·7 by 20,000 kg/cm² pressure. A suspected phase transition in solid helium and a definite one for nitrogen at 65°K are discussed. The Murnaghan theory of finite strain leads to a convenient representation of all the experimental data. A direct determination of the density of solid argon at 77°K made by W. M. HINDS is also reported.

VOL. 1 1956/57

1. INTRODUCTION

THE present work is a continuation and extension of that performed earlier by the author.[1] The experimental measurement of pressure-volume relations of solidified gases at low temperature is of especial interest to solid state physics because these are among the simplest of dielectric solids. Furthermore it is only at low temperature that theoretical calculations of densities are feasible. For most of the cases considered, 4°K is essentially absolute zero, so that thermal effects should be unimportant. The measurements for solid hydrogen and helium are also of concern to astrophysicists who believe that the planets Jupiter and Saturn may consist nearly entirely of solid hydrogen and helium.[2] The pressures at the centers of these planets (millions of atmospheres) are far beyond the experimentally attainable range, but it is possible to extrapolate experimental measurements made at lower pressures with a fair degree of certainty.

Measurements of relative volume changes have been made up to 20,000 kg/cm² for helium, hydrogen, deuterium, and neon at 4°K, for nitrogen and argon at 65°K, and for argon at 77°K. Inasmuch as no careful theoretical calculations of density as a function of pressure seem to exist for these substances, the present paper is concerned primarily with experimental data with little attempt at theoretical interpretation. It is hoped that these results may serve as a stimulus to further theoretical work.

2. APPARATUS AND TECHNIQUE

The piston displacement method, originally developed by BRIDGMAN,[3] was used in the present work. The new 20,000 kg/cm² apparatus is for the most part simply an enlarged version of the earlier low-temperature equipment[1, 4] so only those features which are different in the present apparatus need be described in detail here. The stainless-steel (Type 304) tension and compression members of the press were once again in the form of hollow concentric cylinders. These were about 30 inches long with wall thickness $\frac{1}{8}$ inch, and were of 2 inch and $1\frac{1}{2}$ inch o.d., respectively. The maximum force transmitted to the compression member by the piston of the hydraulic press was about 7 tons, which corresponds to a pressure of 20,000 kg/cm² over the $\frac{1}{4}$-inch experimental piston tip. The pressure in the hydraulic system was measured and could be maintained constant by means of a dead weight piston gauge to an estimated accuracy of 0·5 per cent. The motion of the upper end of the compression member could be followed as the

* Supported by the Office of Ordnance Research, U.S. Army.

pressure was varied to within less than 0·005 mm by means of a micrometer telescope.

As in the past, the technique was to apply a one-sided compressive force directly to the sample of solidified gas contained in a thick-walled cylinder. If measurements of the piston displacement are to give meaningful values of $\Delta V/V_0$, the solid samples must be fairly plastic. Extrusion experiments reported earlier[1] have shown that hydrogen, deuterium, and neon at 4°K, and argon at liquid-nitrogen temperatures, are quite plastic, at least at low pressures. Departures from hydrostatic pressure within the samples will not be large provided that the shear strength of the solid is only a small fraction of the maximum applied pressure.

In the earlier work by the author to 4000 kg/cm²[1] and by SWENSON to 10,000 atm.,[4] the piston tips and sample holders were both made from a heat-treated tool steel. It was found early in the present investigation that cylinders (o.d. 1 inch, i.d. ¼ inch) made from this material containing solid hydrogen burst at pressures of the order of 10,000 kg/cm². The fractures appeared to be extremely brittle. In order to reach 20,000 kg/cm², it was found necessary for the cylinders to use a low-carbon nickel steel, which was more ductile at the low temperature. The best results were obtained with a case-hardenable alloy (AISI 4615) of composition C 0·13–0·18 per cent, Mn 0·45–0·65 per cent, Si 0·20–0·35 per cent, Ni 1·65–2·00 per cent, Mo 0·20–0·30 per cent.[5] The cylinders were heat treated for toughness, but not carburized as there was no advantage in hardening the surface. The piston tips were still made of the hardened Ketos steel to avoid plastic deformation. There was no evidence that the thermal expansivity of these two steels differed by an amount sufficient to spoil the machined fit.

The cylinders were constructed with one closed end. In spite of the greater difficulty in fabrication, this design was essential in the present work because of the need for sealing washers around the moving piston. The open-ended design with two pistons used successfully by SWENSON[4] in his measurements on the alkali metals would double the packing difficulties which are already quite severe with one piston.

The samples were prepared by allowing the gas to condense into the sample holder at low temperature through a 0·035-inch stainless-steel capillary. All except helium at 4°K and nitrogen at 65°K could be frozen at atmospheric pressure. The gases were prevented from leaking out by means of a greased piece of thin brass foil clamped tightly over the top of the cylinder. The brass foil gave a better seal than did the Glassine envelope window used before,[1] and was readily broken by the piston when pressure was applied. To prevent blockage during the filling, a current of 0·5–1·0 ampere was passed through the capillary. The condensation pressure varied between atmospheric and 30 psig (except for He and N_2) which was in all cases above the triple point pressure, so the samples must have passed through the liquid state before freezing.

The preparation of the He samples is worth describing in detail. At 4·2°K the freezing pressure of helium is 143 kg/cm²[6] or 2,040 psi so that at least this much pressure must be applied to the liquid in the sample holder in order to produce solid. As was discovered in the earlier work[1] the most satisfactory packing material at 4°K is metallic potassium. For maximum efficiency in sealing, the unsupported area principle of BRIDGMAN[7] must be utilized. At 4°K potassium flows plastically at a pressure of a few hundred atmospheres; below this pressure it will not seal. A number of attempts were made to solidify liquid helium in the sample holder merely by breaking through the brass seal with the high-pressure piston but the liquid always squirted out before the packing became effective.

Fig. 1 shows the scheme that proved successful. The packing was placed beneath the seal on the top of the sample holder. The seal was made from a disk of copper 0·010 inch thick which had been annealed before being forced into a conical shape. This heavier seal was needed to withstand the high gas pressure in the sample holder. The capillary ran from the sample holder out to room temperature and into a heavy walled cylinder of several cm³ volume which served as a "pressure booster", and thence to a full helium cylinder. After the apparatus had been assembled, but before it was cooled, helium gas at about 2000 psi was admitted to the sample holder. This served to compact the potassium washer while it was soft. When the apparatus was cooled to 4°K the pressure in the capillary dropped several hundred psi as liquid condensed in the sample holder. Then the capillary was electrically heated (to prevent blocking) and the full helium tank pressure of 2500 psi was admitted. The pressure booster was immersed in liquid nitrogen, the valve between it and the helium cylinder closed, and the pressure booster then transferred to a bath of boiling water. When equilibrium had been attained the pressure in the sample holder was estimated to be 6500 psi, well above the solidification pressure. The heating current was then cut off to allow the helium in the capillary to solidify.

Two pumps connected in parallel were used in the hydraulic system so that the pressure could be raised rapidly. When only one pump was used, the sample was lost between strokes after the copper seal had been broken but before the packing sealed. With two pumps worked by hand 180° out of phase this difficulty was avoided. Once the pressure had successfully been raised to a few thousand atmospheres the sample was never subsequently lost. The friction in the packing was sufficiently large so that the oil pressure could be dropped to zero without loss of the sample.

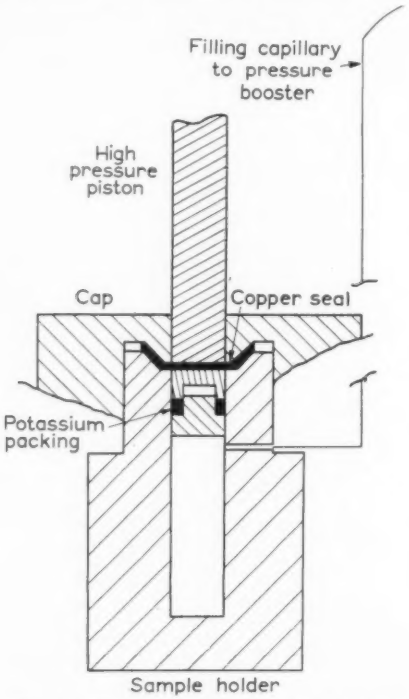


FIG. 1. The solid-helium sample holder.

The same technique as described above was used to produce samples of solid nitrogen at 65°K, for which the solidification pressure is approximately 100 kg/cm².[8]

The low-temperature bath liquid was contained in a conventional all-metal Dewar of five litres' capacity. The surrounding liquid-nitrogen jacket consisted of a commercial (Hofman Laboratories, Inc., Newark, N.J.) 15-litre stainless-steel Dewar. With the pressure apparatus in position, the evaporation rate of helium from the inner Dewar was about one litre of liquid per hour, so that, after losses from initial cooling of the apparatus from 77° to 4°, runs lasting three hours could be made. For runs at 77° and 65° the inner Dewar was filled with liquid nitrogen. The temperature of 65° was obtained by pumping vapor from the bath to 125 to 130 mm of mercury, and could be held constant to ±0·5°K by manual adjustment of a valve in the pumping line.

The technique for taking the data was exactly as previously described.[1, 4] The pressure was rapidly run up to maximum value and released two or three times to insure compaction of the sample. When reproducible telescope readings at maximum pressure were obtained the actual recording of data was started. The pressure was monotonically decreased to zero by removing weights from the dead weight gauge pan in ten to fifteen steps and telescope readings recorded for each value of the pressure. Sufficient time was allowed between readings to maintain isothermal conditions in the sample and to make sure that there was no drift in the telescope readings. At zero pressure the position of a fiducial mark on the observation window was recorded, the telescope refocussed on the compression member, and a set of readings taken with pressure increasing to the maximum value. The initial and final telescope readings for such a run nearly always agreed closely; if they did not the data were rejected. Usually a second run, and occasionally a third, were made before the apparatus was disassembled.

VOL. 1 1956/57

The three important corrections for this type of measurement are:

1. Friction
2. Distortion of the apparatus under pressure
3. Dilation of the sample holder volume from internal pressure.

Friction is the principal factor limiting the accuracy of the results. It is particularly severe at 4°K, and in some of the runs the separation between points of constant telescope reading for increasing and decreasing pressure was as much as 3000 kg/cm², or 15 per cent of the maximum pressure. Most of the friction resides in the sample itself and in the potassium packing; a relatively small amount is contributed by the leather packing in the hydraulic press at room temperature. BRIDGMAN[3] has discussed such friction in connection with his room temperature compressibility measurements and concludes that the most realistic correction is an average between the curves for decreasing and increasing pressure at constant displacement. This was the practice adopted here. Near the two ends of the curve one can no longer interpolate halfway between the smooth curves connecting the experimental points, but must extrapolate parallel to one or the other. See Fig. 2 below. One can have considerable confidence in this procedure if the two curves are parallel, or nearly so, as they were always observed to be. This shows that the friction is nearly constant. In the worst cases the half-height of the hysteresis loop of displacement against pressure was 15 to 20 per cent of the total compression of the sample, so that if the curves are parallel to within 15 per cent to 20 per cent as was always the case, an uncertainty of about 3 per cent is introduced into the final values of $\Delta V/V_0$.

The second correction is for the distortion of the apparatus itself under pressure. This must be evaluated with each sample because the compression of the potassium packing varies from run to run depending upon the thickness of the washer. The solidified gas sample is allowed to escape by temporarily warming up the apparatus, and a hysteresis loop then obtained for decreasing and increasing pressure with the piston pressing on the bottom of the sample holder. A reading of the fiducial mark on the window is also taken so that the sample length may be calculated. The friction in this "blank run" is of the order of 500 kg/cm², much of it being contributed by the packing. Correction is made in the same manner as when the sample is present. Fig. 2 shows plots of the actual data taken in one of the runs with solid helium.

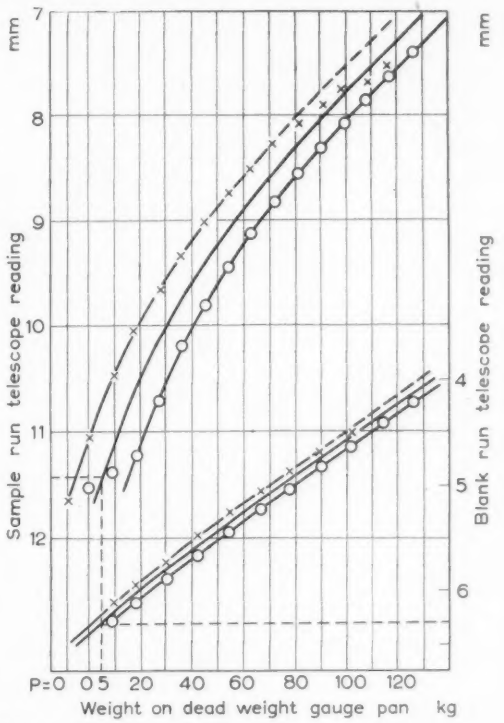


FIG. 2. Experimental data for a typical helium compressibility run, 7 March 1956. The blank run, the friction correction, and the method for obtaining the sample length are shown.

O = Increasing pressure. X = Decreasing pressure.

The ratio of the blank run displacement to the total displacement with a sample ranged from about 0·3 for the longest H_2, D_2, and He samples up to about 0·6 for short samples. Since this correction can be determined considerably more accurately than can the sample compression itself, the uncertainty introduced into the final results is small.

The length of the sample at any desired pressure could easily be calculated from the telescope readings and from the position of the fiducial mark on the observation window. Whenever the sample was expelled before a blank run, the telescope had to be removed. Upon being replaced it was focussed on the fiducial mark with as nearly as possible the same reading each time to eliminate errors resulting from the direction of travel of the telescope not

being exactly parallel to the length of the compression member. It is believed that the sample length in each case is known to better than $\pm 0{\cdot}1$ mm, or one per cent for average length samples.

From the actual telescope readings this length can be calculated considerably more accurately than 1 per cent, but in some cases with the 4°K runs the liquid helium bath was exhausted before the blank run could be carried out. This means that the temperature inside the cryostat was between 4° and 77° (no higher because the liquid nitrogen level in the outer Dewar was maintained constant throughout). SWENSON[4] has found that the compressibility of potassium and the YOUNG's modulus for stainless steel, which together determine the magnitude of the blank run correction, differ by an extremely small amount between 77° and 4°. The most serious effect of not maintaining the temperature of 4° during the blank run is associated with the determination of sample length. Recent measurements by SWENSON[9] of the total thermal contraction of Type 304 stainless steel at low temperature show that the tension and compression members should be shorter by roughly 0·1 mm at 4° than at 77° which introduces an error of less than one per cent into the calculated sample length. In view of the larger uncertainty introduced by the friction, it did not seem appropriate to expend the additional several litres of liquid helium per run which would have been required to carry out the blank run at 4°K.

The third correction, that for sample holder dilation, had to be determined experimentally. It was not feasible to calculate it from elasticity theory because of both the complicated geometry and the uncertainty of the values of the elastic constants for the heat-treated steel at low temperature. The sample holders were filled with varying amounts of indium metal and the apparent compressibility of the indium measured at 77°K in the same manner as for the solidified gases. The difference between the apparent compression and SWENSON's[10] values for indium at 77° (measured only to 10,000 kg/cm² but assumed to be linear with pressure to 20,000 kg/cm²) was attributed to sample holder dilation. The correction was found to be small in all cases, and the difference between 77° and 4° values should be completely negligible. For $\frac{1}{4}$-inch i.d. sample holders the correction amounted to 1·2 per cent of the base volume for long samples, and fell off almost linearly with length for samples shorter than 11 mm. For the $\frac{5}{16}$-inch i.d. holders the correction was 1·0 per cent at the maximum pressure (12,800 kg/cm²) for long samples, and again fell off linearly with sample length for shorter samples. The uncertainty in this determination is quite high (20 per cent) owing to the smallness of the effect, but the error resulting in the final data is small.

When all the sources of error are considered, it appears that the final experimental data for the relative volume decrease are trustworthy to about 5 per cent. The variance among the several runs made with each substance is about of this order of magnitude.

In most cases runs were made in both the $\frac{1}{4}$-inch and $\frac{5}{16}$-inch sample holders, and the sample lengths varied by a factor of two or more. No trends of the magnitude of the relative compression with either sample diameter or sample length were ever observed. This gives one confidence that hydrostatic pressure was transmitted satisfactorily by all the samples. In addition, some measurements were made on solid argon in a $\frac{5}{8}$-inch i.d. sample holder.

Since small amounts of impurity should have very little effect on the observed compressibilities, no special effort was made to obtain gases of high purity. All gases are believed to have been at least 99·0 per cent pure, which is more than adequate in view of the other limitations.

3. CALCULATION OF THE RESULTS

On account of the large and rapidly changing compressibilities of the solidified gases at low pressure, it was impossible to obtain satisfactory data for $\Delta V/V_0$ at pressures below about 2000 kg/cm². It will be seen from Fig. 2 that the friction correction can be made reliably only at higher pressure, and that extrapolating the corrected (central) curve, which is easy at the high-pressure end, is unsafe at low pressure because of the large curvatures. Therefore the base volume has been taken to be that occupied by the sample at $P = 2043$ kg/cm², corresponding to a weight of 5 kg on the dead weight gauge pan. This point is indicated by the dotted lines in Fig. 2.

Values of the relative volume decreases $\Delta V'/V_0'$ refered to the 2043 kg/cm² base volume are calculated directly from smoothed values of the observed compression of the sample diminished by the blank run correction and the sample holder dilation. The assumption that $\Delta x'/x_0' = \Delta V'/V_0'$, where x' refers to the sample length, is justified only so long as the samples are plastic and are transmitters of hydrostatic pressure, as is presumed throughout.

To obtain more meaningful data it is necessary to extrapolate the averaged and smoothed observed values of $\Delta V'/V_0'$ to give $\Delta V/V_0$, where V_0 is the volume at zero (atmospheric) pressure except as noted below. Where previous compressibility data at lower pressure exist, these can be used as the basis of extrapolation, and an estimate of the consistency of the new results obtained from the smoothness of the fit at $P = 2043$ kg/cm². A second method is to represent the $\Delta V'/V_0'$ data analytically over the range $P = 2043$ kg/cm² to $P = 20{,}000$ kg/cm², and then to extrapolate algebraically to $P = 0$.

VOL. 1 1956/5

It has been found by BIRCH[11] that BRIDGMAN'S compressibility data for a large number of substances can be represented very closely in terms of MURNAGHAN'S theory of finite strain. BIRCH obtains an expression for pressure in terms of the compression ratio V_0/V from the assumption that the elastic energy of the compressed solid can be represented in terms of the square (which alone is retained in ordinary infinitesimal elasticity theory) and the cube of the finite strain. He assumes $\Psi = af^2+bf^3$ where Ψ is the elastic strain energy and a and b are constants depending only upon the temperature. The finite compressional strain f for hydrostatic pressure, which is always positive, is found from MURNAGHAN'S theory to satisfy the relation $V_0/V = (1+2f)^{3/2}$. One then obtains

$$P = \frac{3}{2\beta_0}[y^7-y^5][1-\xi(y^2-1)] \qquad (1)$$

where $y = (V_0'/V)^{1/3}$, β_0 is the isothermal compressibility $-(1/V_0)(\partial V/\partial P)_T$ evaluated at $P = 0$, and $\xi = -3b/4a$. We also have the relation $3/2\beta_0 = \alpha/3V_0$. In practice it is not convenient to evaluate the parameters a and b directly so that ξ must be found by trial and error. Equation (1) has been used successfully at low temperature by the author to represent the earlier data for solidified gases, and by SWENSON[4] for the alkali metals.

When equation (1) is to be used to represent data where V_0 is the volume at some pressure other than zero, it should be modified as follows:

$$P-P_0 = \frac{3}{2\beta_0'}[y'^7-y'^5][1-\xi'(y'^2-1)] \qquad (2)$$

where $y' = (V'_0/V)^{1/3}$, P_0 is the pressure at which $V = V'_0$, $\beta'_0 = -(1/V'_0)(\partial V/\partial P)_T$ evaluated at $P = P_0$, and ξ' is an adjustable constant.

The procedure for fitting experimental data to either equation (1) or (2) is to tabulate values of the function $f(y, \xi) = (y^7-y^5)(1-\xi(y^2-1))$ against pressure for different values of ξ and to choose that value of ξ which gives the best straight line. β_0 is then readily calculated from the slope of the straight line.

Equation (2) can be used to extrapolate $\Delta V'/V'_0$ data to give $\Delta V/V_0$ by finding the value of y' (< 1) which makes $P = 0$. Then $y' = (V'_0/V_0)^{1/3}$ and $V/V_0 = (V/V'_0)(V'_0/V_0)$, from which values of $\Delta V/V_0$ can be computed. This method of extrapolation works well for substances whose compressibilities do not change too rapidly with pressure, but it could not be used in the cases of helium, hydrogen, or deuterium, for which half or more of the total compression in 20,000 kg/cm² comes in the first 2000 kg/cm². For these three cases, other experimental data must be used as the basis for the extrapolation.

Once values of $\Delta V/V_0$ have been obtained, one can fit them by trial and error to equation (1). In all seven cases considered in the present investigation, the maximum deviation of the observed pressure from that given by (1) was less than the estimated experimental error.

The most convenient way to calculate the compressibility as a function of pressure is to differentiate equation (1) and use the values of ξ and β_0 found in the original fitting of the relative volume change data. The result is

$$\beta = \frac{2\beta_0}{y^8}[-9\xi y^4+(7+14\xi)y^2-(5+5\xi)]^{-1} \qquad (3)$$

This relation has been used to calculate values of β, from which the *instantaneous* compressibility $\beta_i = -(1/V)(\partial V/\partial P)_T$ can be obtained. The analytical differentiation is more accurate than graphical differentiation, and the uncertainty in the values of β should be roughly 5 per cent, the same as in $\Delta V/V_0$. At low pressure the error may be greater because β is changing rapidly and the fit of the data by equation (1) may not be as good.

Because of the small volume of the sample holders and uncertain temperature gradients along the filling capillary, the present apparatus is not adapted for the direct measurement of densities of the solidified gas samples. Furthermore, a portion of some of the samples may have been lost by extrusion during the initial compaction. Therefore, where molar volume data are given they have been obtained from other sources and may be subject to additional uncertainty.

4. RESULTS AND DISCUSSION

Table 1 presents a summary of the compressibility determinations that were made.

The values for ξ and $3/2\beta_0$ from equation (1) for the seven cases considered are presented in Table 2.

BIRCH[11] has found that many of BRIDGMAN'S results up to $P = 100{,}000$ kg/cm² are quite closely represented by equation (1) with $\xi = 0$. In addition, SWENSON[4] found that sodium, potassium, and rubidium at 4°K satisfy the same relation at pressures up to 10,000 kg/cm². The elastic energy under hydrostatic compression of these substances evidently depends only upon the square of the strain. It appears that the solidified gases behave in a more complicated manner in that all require negative values of ξ. A negative value of ξ means that the constant b in BIRCH'S equation for the

Table 1. Summary of compressibility runs. x_{max} and x_{min} represent the lengths of the longest and shortest samples used, measured in each case at $P = 2043$ kg/cm².

Substance	Temp. °K	¼-inch Holder			5/16-inch Holder			0·626-inch Holder		
		Runs	x_{max} mm	x_{min} mm	Runs	x_{max} mm	x_{min} mm	Runs	x_{max} mm	x_{min} mm
Helium	4	5	12·7	5·0	0	—	—	0	—	—
Hydrogen	4	3	11·5	6·6	0	—	—	0	—	—
Deuterium	4	3	15·6	12·9	1	14·5	—	0	—	—
Neon	4	3	14·4	5·7	1	17·6	—	0	—	—
Nitrogen	65	4	16·4	8·5	1	10·3	—	0	—	—
Argon	65	2	17·2	10·7	2	17·4	17·3	2	21·0	15·7
Argon	77	6	21·5	8·5	1	13·5	—	2	21·3	21·2

Table 2. Values of $3/2\beta_0$ and ξ in equation (1) which give best fit to experimental compression data.

Substance	Temp °K	ξ	$10^4 \times 3/2\beta_0$ kg/cm²	Remarks
Helium	4	$-2·3 \pm 0·1$	0·131	$P_0 = 143$ kg/cm² at melting curve
Hydrogen	4	$-1·9 \pm 0·1$	0·306	
Deuterium	4	$-1·6 \pm 0·2$	0·496	
Neon	4	$-4·0 \pm 0·5$	1·52	
Nitrogen	65	$-2·3 \pm 0·1$	1·91	$P < 10{,}200$ kg/cm². $P_0 = 100$ kg/cm
		$-0·7 \pm 0·1$	(2·86)	$P > 10{,}200$ kg/cm². $P_0 = -1380$ kg/cm²
				Phase transition at $10{,}200 \pm 800$ kg/cm²
Argon	65	$-4·9 \pm 0·5$	2·40	
Argon	77	$-6·5 \pm 0·5$	1·63	

strain energy must be positive, since a is necessarily positive, which means that these substances are too "hard" when compared with "normal" substances which have $\xi = 0$. Their compressibilities decrease more rapidly with pressure than would be the case for a "normal" substance. However, it should be emphasized that equation (1) has no direct theoretical significance in terms of an interatomic force model and is thus merely a convenient analytical representation for the experimental data.

We now pass to a detailed presentation of the data and a separate discussion for each of the solidified gases. Table 3 presents all the data for 4°K, while the results for nitrogen and argon are given in Table 4. In both tables the original experimental data for the volume change relative to the volume at $P = 2043$ kg/cm² are given, since these results are not subject to possible errors resulting from extrapolation to zero pressure. The instantaneous compressibilities β_i are believed to be more meaningful for comparison than $\beta = -(1/V_0)(\partial V/\partial P)_T$.

Helium. Oil-free grade "A" commercial helium such as can be handled by the Collins liquefier without a cold trap was used. Because of the elaborate seal necessary to hold the samples, the friction for these runs tended to be greater than for the others, so the overall accuracy of the results may be poorer. Unlike the other solidified gases which have been shown[1] by extrusion measurements to be fairly plastic, there is no direct evidence on how effective a transmitter of hydrostatic pressure solid helium is. On account of the sealing difficulties, it has not yet been possible to carry out extrusion experiments, or even to prepare samples in the 5/16-inch sample holder. However, the ¼-inch

Table 3. Compressions at 4°K.

Pressure kg/cm²	Helium				Hydrogen				Deuterium				Neon			
	$\frac{\Delta V'}{V_0'}$	$\frac{V}{V_0}$	M.V. cm³	$10^5\,\beta_i$ cm²/kg	$\frac{\Delta V'}{V_0'}$	$\frac{V}{V_0}$	M.V. cm³	$10^5\,\beta_i$ cm²/kg	$\frac{\Delta V'}{V_0'}$	$\frac{V}{V_0}$	M.V. cm³	$10^5\,\beta_i$ cm²/kg	$\frac{\Delta V'}{V_0'}$	$\frac{V}{V_0}$	M.V. c.m³	$10^5\,\beta_i$ cm²/kg
P_0 or 0	—	1·000	17·1	115	—	1·000	22·65	49	—	1·000	19·56	30	—	1·000	14·0	10
200	—	·924	15·8	77	—	·928	21·0	31	—	·948	18·5	22	—	·982	13·7	—
400	—	·846	14·5	41	—	·883	20·0	23	—	·911	17·8	18	—	·966	13·5	—
600	—	·792	13·5	29	—	·847	19·2	19	—	·882	17·2	15	—	·952	13·3	—
1000	—	·722	12·4	19	—	·794	18·0	14	—	·840	16·4	12	—	·927	13·0	5·3
2043	0	·625	10·7	9·8	0	·711	16·1	8·1	0	·764	14·9	7·3	0	·882	12·3	3·8
3000	·079	·579	9·9	7·0	·065	·667	15·1	6·1	·061	·715	14·0	5·5	·032	·858	12·0	—
4000	·134	·544	9·3	5·5	·110	·632	14·3	4·8	·103	·680	13·3	4·5	·055	·832	11·6	2·6
6000	·205	·497	8·5	3·8	·181	·583	13·2	3·5	·173	·632	12·4	3·3	·100	·797	11·1	2·0
8000	·256	·466	8·0	2·9	·227	·549	12·4	2·7	·215	·597	11·7	2·6	·130	·770	10·8	1·7
10,000	·295	·440	7·5	2·4	·266	·523	11·8	2·3	·253	·571	11·2	2·2	·156	·747	10·4	1·41
12,000	·325	·421	7·2	2·0	·297	·500	11·3	1·9	·281	·549	10·7	1·9	·178	·728	10·2	1·23
16,000	·370	·392	6·7	1·5	·343	·467	10·6	1·4	·325	·514	10·0	1·4	·213	·696	9·7	·98
20,000	·403	·373	6·4	1·3	·376	·445	10·1	1·2	·358	·485	9·5	1·2	·240	·669	9·4	·81

samples used varied in length by a factor of 2½ with no systematic trend in the observed compression. Therefore it is felt quite strongly that solid helium is sufficiently plastic so that this method of measurement is valid.

The only means at hand for extrapolating the $\Delta V'/V'_0$ data to the melting pressure at 4°K was by means of the isochores given by Dugdale and Simon.[6] They made no direct compressibility determinations but were able to calculate thermodynamically molar volumes as functions of pressure and temperature from their measurements of the specific heat at constant volume at various densities and knowledge of the melting parameters. Values of molar volume at 4·2K were read off their diagram for different pressures between the melting pressure and 2200 kg/cm². Their value of V at 2043 kg/cm² (obtained by interpolation) was used for the basis of the extrapolation. It is quite gratifying that their curve for V below 2043 kg/cm² fits smoothly onto ours for higher pressure when the extrapolation is made in this manner. A good fit to the Murnaghan relation, equation (1), is obtained right across the junction of the two sets of data. The compressibilities as calculated from (3) for pressures below 2043 kg/cm² agree well with those for 0°K given by Dugdale and Simon.

Solid helium is seen to be the most compressible of all solids, as has long been suspected. The main reason for this is its large zero-point energy which causes its volume at zero pressure to be much larger than it otherwise would be. This is the cause of the very rapid decrease of compressibility with pressure—by a factor of ten in the first 2000 kg/cm², and by almost another factor of ten at $P = 20{,}000$ kg/cm². At the higher pressures where the elastic energy of compression is large compared to the zero-point energy, the compressibility is almost exactly the same as for hydrogen and deuterium, and less than double that for neon and argon (at 77°K). If one takes for the zero-point energy its value on the melting curve at 0°K of 50 cal/mole,[12] one finds that the work of compression becomes of the same order of magnitude at a pressure of about 1000 kg/cm². The very rapid decrease of compressibility with pressure appears to slow down near $P = 6000$ kg/cm², at which point the total work of compression is roughly six times the zero point energy. At 20,000 kg/cm² the elastic energy is seventeen times the initial zero-point energy. This calculation is of only qualitative significance because the Debye θ and therefore the zero-point energy increase sharply with decreasing volume,[6] but it would appear that the compressibility at the top of the pressure range

should not be greatly influenced by the zero-point energy.

When the helium data were first analysed there was a strong suspicion that there was a phase change in the solid at a pressure of 10,000 to 11,000 kg/cm². It appeared that two values of the constants ξ and β_0 of equation (1) were needed to fit the data below and above this pressure. Subsequent examination in which equation (1) was applied to each of the runs separately showed that not all of them exhibited signs of a transition. The friction is so large that a sudden change in the compressibility or even a small volume discontinuity would be smeared over a wide pressure range and practically unobservable. The known transition in solid nitrogen, to be discussed below, was barely discernible in the data for that substance and friction was less severe in the nitrogen runs than with the helium. The present experimental data are not sufficient to confirm or deny the existence of a transition in the neighborhood of 10,000 kg/cm². Because of this doubt as to the reality of the transition, only one value of ξ was used to represent the data. The fit at $P = 20{,}000$ kg/cm² is still within experimental error. Further measurements are contemplated.

The question is of particular interest because of the first order transition in the solid found at higher temperature by DUGDALE and SIMON.[6] Their transition line intersects the melting curve at $T = 14{\cdot}9°$K, but the initial slope is definitely towards higher temperature with higher pressure. It is unlikely that there could be a third phase at the lower temperature because an inert substance like helium must crystallize in a close-packed structure and there are only the two possibilities. DUGDALE and SIMON believe that their transition is from close-packed hexagonal to close-packed cubic. Therefore, if this transition is to occur at 4°, the equilibrium line between the two phases must bend back on itself. At the melting curve the volume change in the transition (calculated from the Clapeyron equation and measured thermal quantities) was found to be only 4×10^{-4} cm³ or of the order of one part in 30,000. Even in the absence of friction this could never be observed as a volume discontinuity with a piston displacement apparatus. The best one could hope for would be a break in the compressibility as calculated from the $\Delta V/V_0$ data.

Theoretical calculations by BARRON and DOMB[13] on the relative stability of the two close-packed lattices have indicated that there should be a first-order transition at absolute zero from close-packed hexagonal to close-packed cubic at high pressure. In the case of helium they find that the transition should start out from the melting curve at 15°K as observed by DUGDALE and SIMON, but that at higher pressure it should reverse its direction and go to lower temperature. BARRON and DOMB conclude that the energy difference between the two lattices is always very small (of the order of 0·01 per cent), so that the volume change at the transition should be correspondingly small. At 4°K their calculations indicate a transition at a molar volume of 4·5 cm³; the suspected effect in the present data occurs at 7·4 cm³/mole. The discrepancy is large, but the computations are subject to error because of the extremely small energy difference involved. If the data of Table 3 are extrapolated using equation (1), a molar volume of 4·5 cm³ is found to correspond to a pressure of 75,000 kg/cm². Conclusive experimental verification of this effect must await a means for substantially reducing friction in solidified gas samples at 4°K, and possibly also extension of the measurements to higher pressure.

Hydrogen and Deuterium. Standard commercial hydrogen was used. The deuterium was supplied by the Stuart Oxygen Company with a purity of better than 99·5 per cent. The present results confirm the earlier ones[1] very closely. It was necessary to use the value of $\Delta V/V_0$ at $P = 2043$ kg/cm² that had been obtained before[1] to carry out the extrapolation of the present $\Delta V/V'_0$ data to zero pressure. When this is done the new $\Delta V/V_0$ data agree almost perfectly with the old over the entire pressure range of the latter, which was 0 to 10,000 kg/cm² for both hydrogen and deuterium. The constants in equation (1) needed to fit the new data up to 20,000 kg/cm² are the same as before, so that the new results could have been obtained simply by extrapolating the old.

Only three runs, all with $\frac{1}{4}$-inch samples, were deemed necessary for hydrogen because of the excellent agreement both among themselves and with the earlier results. There is evidence that the plasticity of solid hydrogen continues to be very high at high pressure both from the mode of bursting of the tool-steel sample holders used at the beginning of this investigation and the extremely rapid extrusion and loss of sample that occur in the absence of packing.

The values of the molar volumes of hydrogen and deuterium at zero pressure used in Table 3 were obtained by MEGAW.[14]

Zero point energy again plays a predominant role in the very high initial compressibilities of H_2 and D_2, as well as in the rapid decrease of compressibility at high pressure. At high pressure the molar volumes of H_2 and D_2 approach each other but they still differ by 6 per cent at 20,000 kg/cm². This means that the zero-point energies still play

VOL. 1 1956/5

a role as otherwise the molar volumes should be the same. Of course the zero-point energy increases to some extent with pressure, but not so fast as the compressional energy so that the relative importance of the zero-point energy diminishes as the pressure is raised.

It is interesting to note that at $P = 20{,}000$ kg/cm² the energies of compression of H_2 and D_2 are both of the order of 30 per cent larger than for helium in spite of the larger compression of the latter. This comes about because of the larger initial molar volumes of hydrogen and deuterium and because of the greater relative decrease of the compressibility of helium at pressures near the top of the range.

Neon. The neon used was the Linde Air Co. research grade from a small cylinder. The extrapolation to zero pressure was made with the aid of the Murnaghan formula, equation (1). The compressibility changes slowly enough at low pressure so that this method is reliable. The $\Delta V/V_0$ data

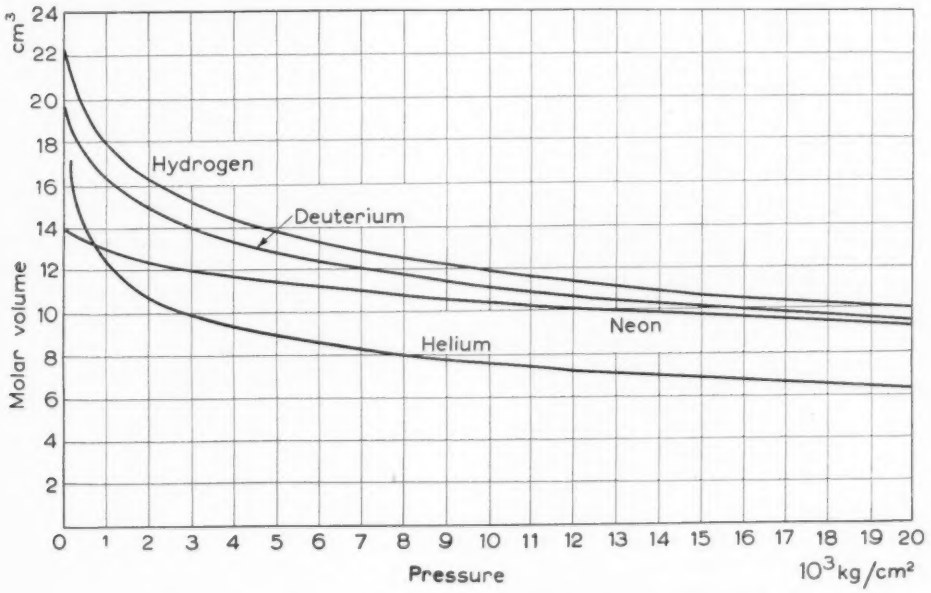


FIG. 3. The molar volumes of helium, hydrogen, deuterium and neon at 4°K.

agree satisfactorily with the older results[1] up to $P = 4000$ kg/cm². The molar volumes have been calculated from a value for the density of 1·443 gm/cm³ given by PARTINGTON.[15] Except at low pressure the molar volume of neon is larger than that of helium because the Ne atom is larger than the He atom. The reversal below $P = 700$ kg/cm² is a consequence of the helium zero-point energy.

Nitrogen. Air Reduction Co. oil-pumped dry nitrogen was used. The measurements with nitrogen served as a rehearsal for helium since at 65° a pressure of 100 kg/cm² must be applied to liquid nitrogen to solidify it. The extrapolation below $P = 2043$ kg/cm² was made with the aid of equation (1). The compressibility appears to change slowly enough at low pressure so that this should be valid.

The principal interest in the results is the confirmation at higher temperature of the phase transition found in solid nitrogen by SWENSON.[16] All of the runs at 65°K show evidence of a sudden change in compressibility, which, after friction has been eliminated, occurs at $P = 10{,}200 \pm$

Table 4. Compressions of Nitrogen and Argon

Pressure kg/cm²	Nitrogen 65°				Argon 65°			Argon 77°K			
	$\frac{\Delta V'}{V_0'}$	$\frac{V}{V_0}$	M.V. cm³	$10^5 \times \beta_i$ cm²/kg	$\frac{\Delta V'}{V_0'}$	$\frac{V}{V_0}$	$10^5 \times \beta_i$ cm²/kg	$\frac{\Delta V'}{V_0'}$	$\frac{V}{V_0}$	M.V. cm³	$10^5 \times \beta_i$ cm²/kg
P_0 or 0	—	1·000	29·3	7·9	—	1·000	6·2	—	1·000	24·3	9·2
500	—	·970	28·4	6·4	—	·975	4·9	—	·964	23·4	6·0
1000	—	·941	27·6	5·2	—	·954	4·0	—	·939	22·8	4·7
2043	0	·897	26·3	3·9	0	·920	3·0	0	·901	21·9	3·3
4000	·059	·841	24·6	2·8	·046	·877	2·1	·051	·853	20·7	2·2
6000	·105	·802	23·5	2·1	·082	·845	1·63	·089	·820	19·9	1·68
8000	·138	·772	22·6	1·77	·108	·820	1·34	·115	·795	19·3	1·40
10,200	·167	·746	21·9	1·50	·131	·798	1·15	·140	·774	18·8	1·18
10,200	·171	·744	21·8	1·45	—	—	—	—	—	—	—
12,000	·190	·727	21·3	1·32	·148	·783	1·03	·157	·758	18·4	1·05
16,000	·226	·693	20·3	1·10	·180	·755	·84	·189	·730	17·7	·84
19,000	·251	·671	19·5	·98	·200	·737	·74	·208	·714	17·3	·74

800 kg/cm². The best fit to the data appears to be with two values of the Murnaghan constants ξ and $3/2\beta_0$, as is shown in Table 2. The negative value of P_0 corresponding to $P > 10{,}200$ kg/cm² has no physical significance. It was obtained by extrapolating the $f(y, \xi)$ vs. P straight line back to $f(y, \xi) = 0$. One can only guess at the value of ΔV at the transition. Swenson states that it is only 0·21 cm³, or 0·8 per cent at the α–β–γ triple point and that it decreases rapidly with higher pressure along the β–γ equilibrium line. The Murnaghan fit seems to be best with $\Delta V = 0{\cdot}3$ per cent at 65°. Such a small volume discontinuity would be difficult to observe directly even in the complete absence of friction. At lower temperatures Swenson had less friction because he used no packing, but this would not be possible at 65° where the sample is initially liquid.

In the earlier work,[1] a single rough determination of β was made for solid nitrogen at 4°K. The agreement with the present results is poor both because of the large difference in temperature and the fact that solid N_2 appears to be somewhat brittle at 4°. No sign of Swenson's α–γ transition was noted then because it would have occurred very close to the top of the pressure range (at 4°, $P = 3500$ kg/cm²) where the friction correction is uncertain.

There is considerable diversity in published values for the density of solid nitrogen [8, 16] so that the estimate of molar volume at 65° at the melting curve of $29{\cdot}3 \pm 0{\cdot}3$ cm³ is less accurate than the other molar volumes used. This value was obtained from extrapolation of Swenson's[16] data at 39·5°K and 53°K. The β–γ phase transition thus appears to occur at $V = 21{\cdot}9 \pm 0{\cdot}4$ cm³/mole at 65°K.

The values of β_i given in Table 4 were calculated by means of equation (3). These are less accurate in the neighborhood of the transition than elsewhere because of the rounding of the discontinuity in β that is inevitable when the friction correction is made. According to Swenson, the compressibility decreases by about 20 per cent as the transition is crossed in the direction of increasing pressure. Table 4 shows a discontinuous decrease of less than 5 per cent which is without doubt too small. Compressibilities determined graphically from the slope of the V/V_0 vs. P curve above and below the transition differ by 12 per cent.

Argon. Commercial argon was used, so this was probably the least pure of all the gases, of the order of 99·0 per cent. Determinations of compression were made at both 65° and 77° because the compressibility at low pressure appears to depend strongly upon temperature in this temperature range. At both temperatures extrapolation to $P = 0$ was made by means of the Murnaghan

VOL. 1 1956/57

relation. The resulting $\Delta V/V_0$ data agree satisfactorily with the earlier work[1] as do the values of ξ and $3/2\beta_0$.

BARKER and DOBBS[17] have recently completed measurements of the adiabatic compressibility of solid argon by means of an ultrasonic interferometer. They measured both longitudinal and transverse wave velocities in polycrystalline samples of solid argon which had been slowly condensed from the gaseous phase below the triple point pressure. These measurements improved upon the earlier results by the same investigators[18] in which only longitudinal velocities had been measured. This method yields values of the compressibility only at zero pressure. BARKER and DOBBS combine their values of the adiabatic compressibility with other data to calculate the initial isothermal compressibility β_0 at different temperatures. At 65°K they obtain $\beta_0 = 5{\cdot}2\times10^{-5}$ cm²/kg, while Table 4 gives $\beta_0 = 6{\cdot}2\times10^{-5}$ cm²/kg. At 77° the values are $6{\cdot}0\times10^{-5}$ cm²/kg and $9{\cdot}2\times10^{-5}$ cm²/kg, respectively. Allowing for an extra 5 per cent error introduced into our values of β_0 by the extrapolation we see that the results at 65° no longer disagree badly, but that there remains a serious discrepancy at 77°.

BARKER and DOBBS believe that the discrepancy may be caused by incomplete compaction of the samples used in the piston displacement measurements. These were prepared by condensation through the liquid state, and were initially quite open in their structure as was shown by the irreversible volume decrease (of the order of 20 per cent) during the initial compaction. However, the reproducibility of the telescope readings, and particularly the fact that the data with maximum pressure 19,000 kg/cm² agree so well with those for which $P_{max} = 4000$ kg/cm²[1] make it highly improbable that the samples are porous. If some pores had remained at 4000 kg/cm² the higher pressure should have been more effective in closing them without affecting the reproducibility of either set of data. If our compressibilities have been too high for this reason, the present results for β should have been lower than the previous ones. This is not the case.

To verify that no large error was introduced through the extrapolation to low pressure, a separate series of runs was made for argon at both 65° and 77° in a large stainless-steel sample holder of inner diameter 0·626 inch. The maximum pressure reached with this holder is 3000 kg/cm², and extrapolation need be made only below $P = 325$ kg/cm². Over the range 325–3000 kg/cm² the agreement with the results in Table 4 was satisfactory. Equation (1) was used to extrapolate over the short range $P = 325$ kg/cm² to $P = 0$. A graphical estimate for β_0, which is subject to large error, yielded values higher than those in Table 4. A more meaningful graphical determination of β at low pressure can be made at $P = 325$ kg/cm² where the friction correction should be reliable. This gives $\beta = 5{\cdot}8\times10^{-5}$ cm²/kg at 65°K and $6{\cdot}4\times10^{-5}$ cm²/kg at 77°. These values can be regarded as lower limits for β_0 so that it would appear that the results of BARKER and DOBBS may be a bit too low.

Another possible source of the discrepancy arises from uncertainty as to the density of solid argon, which needs to be known in the computation of compressibility from the measured sonic velocities. Capt. W. M. HINDS in this laboratory has made a determination of the density at 77°K[19] in order to check the previous values.[17] His method was to allow argon to condense from a 40-litre gas holder at room temperature into a 5 ml Pyrex volumetric flask. The mass of gas in the gas holder before and after condensation was determined from the initial and final pressures, the volume of the gas holder, and the temperature. The volume of the gas holder was determined to 0·1 per cent by direct measurement. The virial coefficients of argon gas[20] were used in the calculation of mass. The Pyrex flask was gradually lowered into a bath of liquid nitrogen so that solid formed in uniform laminae beginning at the bottom of the flask. The rate of growth was of the order of 0·1 mm/min, which is comparable to that obtained by BARKER and DOBBS during the growth of their samples. HINDS' solid was always transparent and uniform in appearance, although great care had to be exercised to prevent the formation of bubbles. It was thus believed to be non-porous. A total of 28 runs were made, some with a starting pressure above the triple point (516·8 mm Hg), and others below. The total pressure drop in the gas holder during condensation was about 100 mm Hg.

The results for the density clustered into two groups, depending upon whether the solid had formed from the liquid, or directly from the gas. The density was found to be $1{\cdot}644\pm0{\cdot}006$ gm/cm³ if the solid had frozen from liquid, and $1{\cdot}627\pm0{\cdot}006$ gm/cm³ for starting pressures below the triple point. In general, the higher the starting pressure, the higher was the observed density. The temperature may be taken to be $77{\cdot}0\pm0{\cdot}5$°K, with slight variations caused by impurities in the nitrogen bath that was used.

For comparison with the densities of BARKER and DOBBS, we should take $\rho = 1{\cdot}627$ gm/cm³, as in their experiment the argon did not pass through the liquid phase. At 77°, they use the value $\rho = 1{\cdot}616$ gm/cm³, so that the agreement at this temperature is certainly satisfactory considering the rather large experimental errors in some of the older density measurements that were used to obtain the above figure. The discrepancy between the current β_0 and that of BARKER and DOBBS is thus not to be explained by an error in density.

Even if our extrapolation below $P = 325$ kg/cm² were incorrect, and the result of BARKER and DOBBS therefore not significantly too low, the values of V/V_0 in Table 4 would be affected very little. The only part of the pressure-volume

relation that would be wrong would be at the extreme lower end. A separate experiment to obtain a precise value of the initial compressibility by direct means would be highly desirable.

Since the samples were all frozen from the liquid, the molar volumes at 77° in Table 4 are based upon $\beta = 1{\cdot}644$ gm/cm³. No molar volumes are given at 65° because of the uncertainty in the density. If BARKER and DOBBS's value is used, one obtains the inconsistent result that $V_{65} > V_{77}$ at $P = 19{,}000$ kg/cm². This could be due to a small error in either the densities or the compressibility data.

Acknowledgements—The author is particularly grateful to Capt. W. M. HINDS who has helped with this work during nearly the whole two years it has been in progress. He assisted in the taking of much of the compressibility data. In addition, he made some of the calculations, and carried out the solid argon density determination. Prof. J. W. BEAMS has given constant encouragement to this project from its inception forward. Mr. PHILIP SOMMER and Mr. FRITZ LINKE of the laboratory machine shop made much of the equipment. Thanks are also due Prof. L. G. HOXTON for helping with the manufacture of the large quantities of liquid helium that were consumed. Dr. C. A. SWENSON was kind enough to make available some of his results ahead of publication.

REFERENCES

1. STEWART J. W. *Phys. Rev.* **97,** 578 (1955).
2. WILDT R. *Planetary Interiors. Handbook of the Solar System*, ed. by G. P. Kuiper, Vol. III, in press.
3. BRIDGMAN P. W. *Proc. Amer. Acad. Arts and Sciences* **74,** 425 (1942).
4. SWENSON C. A. *Phys. Rev.*, **99,** 423 (1955).
5. *Alloy Steel Reference Book*, p. 200 3rd edition, J. T. Ryerson and Son Inc., N.Y. (1951).
6. DUGDALE J. and SIMON F. *Proc. Roy. Soc.* **A218,** 291 (1953).
7. BRIDGMAN P. W. *The Physics of High Pressure* G. Bell and Sons, London (1952).
8. BRIDGMAN P. W. *Proc. Amer. Acad. Arts and Sciences* **70,** 1 (1935).
9. SWENSON C. A. *Rev. Sci. Instrum.* **26,** 1204 (1955).
10. SWENSON C. A. *Phys. Rev.* **100,** 1607 (1955).
11. BIRCH F. *J. Geophys. Research* **56,** 227 (1952).
12. SIMON F. E. and SWENSON C. A. *Nature* **165,** 829 (1950).
13. BARRON T. H. K. and DOMB C. *Proc. Roy. Soc.* **A227,** 447 (1955).
14. MEGAW H. D. *Phil. Mag.* **28,** 129 (1939).
15. PARTINGTON J. R. *Advanced Treatise on Physical Chemistry*, Vol. III, page 149, Longmans Green, London (1952).
16. SWENSON C. A. *J. Chem. Phys.* **23,** 1963 (1955).
17. BARKER J. R. and DOBBS E. R. *Phil. Mag.* **46,** 1069 (1955).
18. BARKER J. R., DOBBS E. R., and JONES G. O. *Phil. Mag.* **44,** 1182 (1953).
19. HINDS W. M. Master's Thesis, University of Virginia, 1956.
20. National Bureau of Standards Circular 564 *Tables of Thermal Properties of Gases*, 1955.

VOL. 1 1956/5

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 159–163.

PHOTOELECTRIC EMISSION FROM THE VALENCE BAND OF CESIUM IODIDE

H. R. PHILIPP AND E. A. TAFT

General Electric Research Laboratory, Box 1088, Schenectady, New York

(*Received* 10 *August* 1956)

Abstract—Photoelectric emission has been obtained both from evaporated films and from single crystals of CsI. For $h\nu < 6$ eV, the yield is relatively low and varies widely from sample to sample. It is ascribed here to direct ionization and to exciton-induced emission from uncontrolled impurities and imperfections. For $h\nu > 6$ eV, the yield rises abruptly to much higher values in a reproducible way for all samples. This effect is attributed to emission from the upper edge of the valence band lying about 6 eV below the vacuum level. Energy distributions of the photoelectrons are consistent with these views. Concurrent optical absorption measurements show detail consistent with band-to-band transitions if the electron affinity of CsI is judged to be less than 0·3 eV.

1. INTRODUCTION

For $h\nu < 6$ eV, alkali iodides do not emit photoelectrons unless they contain imperfections.[1] When F centers are deliberately introduced, two mechanisms contribute to the emission. (1) In the region of optical transparency, F centers may be directly ionized by incident photons. (2) In the first fundamental absorption band, excitons are produced which ionize F centers in a secondary process. In RbI, for example, the direct ionization gives a photoelectric yield of order 10^{-4} electrons per incident quantum when the F center density is about 10^{19} cm^{-3}. The exciton-induced photoelectric yield may rise above 10^{-2} in the first fundamental absorption peak.

There is an additional area in which our understanding is relatively poor. This concerns transitions from the valence band to the conduction band with the resultant production of free electrons and free holes. One expects such a process to begin at a value of $h\nu$ roughly 1 eV larger than that at the first exciton peak.[2] In the past, various peaks in the optical absorption have been tentatively correlated with this ionization limit, but the assignments were not completely satisfying because of possible confusion with exciton processes at higher energy.

This paper discusses new information that bears on this question. The photo-electric emission from relatively pure CsI is low for $h\nu < 6$ eV, consistent with direct ionization and exciton-induced emission arising in a small concentration of defects. Just above 6 eV a new process sets in. The yield rises abruptly above the variable background and reaches values above 10^{-1}. We believe that this is emission from the edge of the valence band, which thus lies about 6·0 eV below the vacuum level in CsI.[3]* Since the bottom of the conduction band is presumably at or below the vacuum level (electron affinity $\geqslant 0$), optical absorption associated with band-to-band transitions must occur for $h\nu$ between 6·0 eV and 5·6 eV, the position of the first exciton peak. Careful examination of concurrent optical absorption data shows a small shoulder which may arise in this way. Thus, we estimate that the electron affinity of CsI is $< 0{\cdot}3$ eV.

2. EXPERIMENTAL TECHNIQUES

Phototubes of three different designs were employed. (1) Cylindrical collectors with central emitters were used to determine the spectral distribution of photoelectric yield at 300°K.[4] Emission measurements were made on evaporated films and on single crystals.†

* A very similar kind of emission occurs in BaO.

† The CsI was obtained in crystalline form from the Larco Nuclear Instrument Company. Pure CsI powder was obtained from Fairmount Chemical Company.

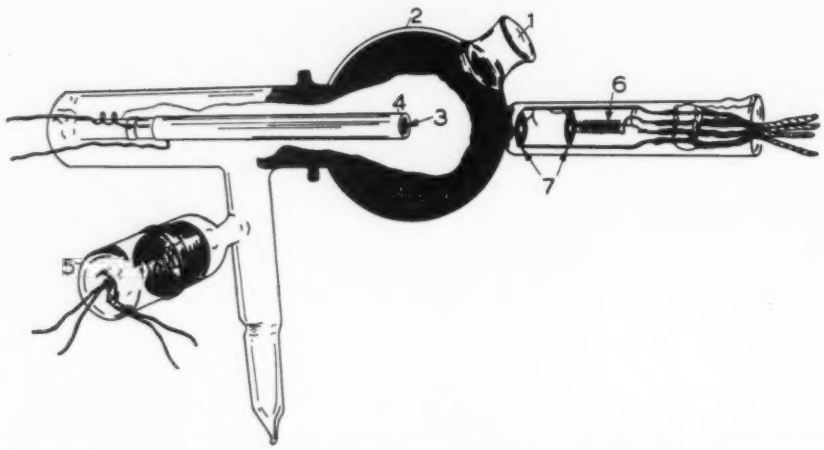


FIG. 1. Cutaway sketch of phototube; (1) 9741 glass bubble window, (2) graphite coated collector sphere 4 inches in diameter, (3) $\frac{3}{8}$ inch glass tube, platinum painted, (4) nickel sleeve insulated from tube by glass beads, (5) ion gauge, (6) evaporating cylinder and helical platinum heater, (7) collimating shields.

(2) Larger phototubes like those used in previous alkali halide work provided spectral distributions at 80°K and at temperatures up to 500°K.[5] (3) A spherical collector with a new type of central emitter was used to measure the energy distribution of emitted electrons in the range 80 to 500°K. A schematic diagram of this tube is shown in Fig. 1.[1, 6]

Optical absorption measurements in the temperature range 80 to 500°K were made on thin films. Material was evaporated from a platinum cylinder heated by a platinum helix. The beam, collimated by $\frac{1}{8}$-inch holes in two molybdenum plates located above the source, was directed at a polished sapphire substrate firmly clamped to a gold-plated copper extension from the coolant reservoir. Interference fringes on the upper collimating plate permitted an estimate of film thickness. Temperatures up to 500°K could be obtained by inserting silicone oil in the reservoir along with a small resistance heater. The light beam was incident on the film at an angle of about 30° from the normal. The phototube windows were thin bubbles of Corning 9741 glass.

3. SPECTRAL DISTRIBUTION OF THE PHOTOELECTRIC YIELD

The spectral distribution of the yield for several samples at 300°K is shown in Fig. 2. Results vary for $h\nu < 6$ eV. Just above 6 eV, the yield rises rapidly to values above 10^{-2} electrons per incident quantum. The quartz monochromator employed in these measurements was not useful for $h\nu > 6{\cdot}6$ eV. Data were extended to 7·4 eV in several cases, using a vacuum monochromator.[7] Yields then rose above 10^{-1}.

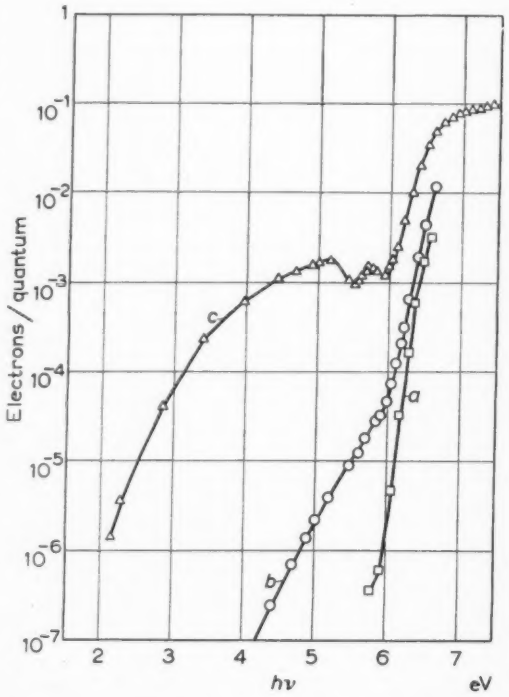


FIG. 2. Spectral distribution of the photoelectric yield for CsI surfaces: (a) thick film, (b) single crystal, (c) thin film evaporated in presence of excess Cs.

Yield curves obtained at different temperatures are shown in Fig. 3. The rapid rise in yield shifted toward lower photon energy with increasing temperature. Curves at all temperatures approached the same high yield above 6·6 eV.

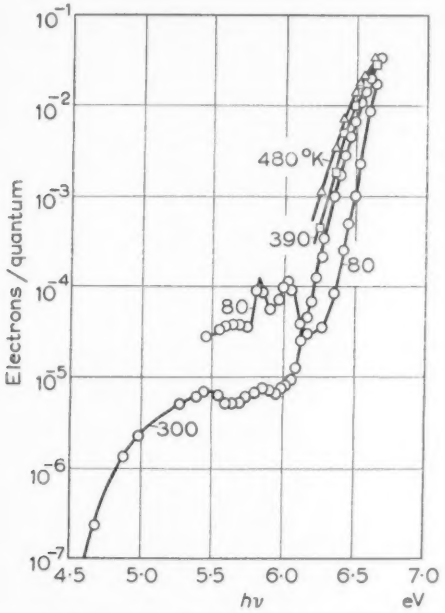


FIG. 3. Spectral distribution of the photoelectric yield for a thin CsI film in the region of fundamental emission at different temperatures.

These results we interpret thus: the variable yield for $h\nu < 5{\cdot}3$ we believe to be direct ionization mainly of unidentified impurities and defects in the CsI. (At 80°K, a low concentration of F centers may be introduced by irradiation in the fundamental absorption bands of this salt; this is not as easily controlled in CsI as in KI or RbI.) Above 5·3 eV fundamental absorption sets in, sharp structure occurring at 5·8 and 6·0 eV when $T = 80$°K. In this region, we expect exciton-induced photo-emission. In some cases, particularly with deliberately introduced F centers, this emission rises to peak values at 5·8 and 6·0 eV, in the manner so common for KI and RbI. In most instances for CsI, however, the exciton-induced emission appears less efficient (by a factor of 2 to 7) than direct ionization, and the yield shows minima at the optical absorption peaks. (Similar effects sometimes occur in KI and RbI, especially when imperfections more complicated than F centers are present.)

In contrast to this extremely variable emission that depends on defect concentration, the yield above 6 eV is quite reproducible in form; it varies only slightly in magnitude among samples. (It is very similar in character to the yield from BaO above 5 eV.) We interpret it as emission from the occupied band of energy states in CsI, which thus lies about 6 eV below the vacuum level.

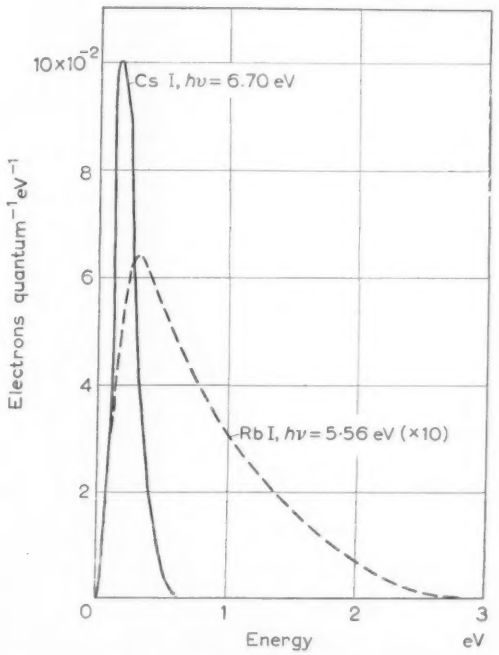


FIG. 4. Energy distribution of the photoelectric yield for CsI in the region of fundamental emission, $h\nu = 6{\cdot}70$ eV. The dashed curve, included for purposes of comparison, shows data on RbI for a value of $h\nu$ near the top of the exciton-induced emission peak. The ordinates of this curve have been multiplied by 10.

4. ENERGY DISTRIBUTION OF THE PHOTOELECTRONS

Fig. 4 shows energy distributions for photoelectrons emitted from CsI at 300°K with $h\nu = 6{\cdot}7$ eV. The photo-electrons were grouped mainly in a band of energies between 0 and 0·6 eV, consistent with emission from a band of levels 6 eV below the vacuum. More than 99 per cent of

the yield thus arose in the occupied band of energy states. These curves are much like those obtained on BaO for $h > 5$ eV.

This distribution is quite different from that for exciton-induced emission, as is illustrated by the broken curve in Fig. 4, which shows the energy distribution of photo-electrons emitted by RbI illuminated in its exciton peak. Both distributions show large numbers of slow electrons, but in the latter case the distribution tapers off much more gradually at higher energies.

5. OPTICAL ABSORPTION

Absorption constants for evaporated films of CsI are shown in Fig. 5. These measurements are

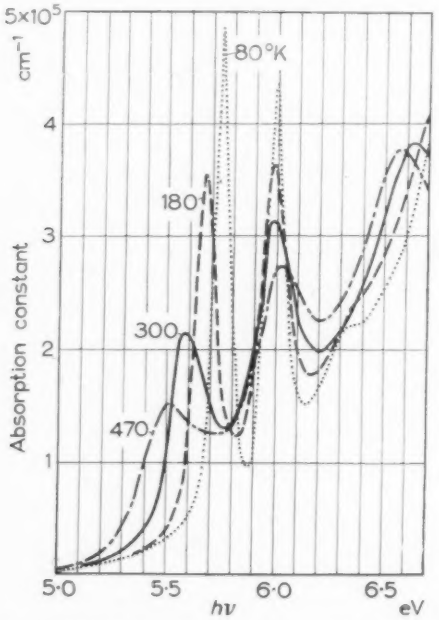


FIG. 5. Optical absorption of an annealed film of CsI. Absolute values of absorption constant are estimated and may be as much as two-fold higher.

in general agreement with those reported previously in the literature,[8] but show more detail. The curves are for films annealed at 200°C for one hour. The films as evaporated on the room temperature substrate show all absorption peaks shifted several hundredths of a volt toward higher energy compared with the peaks of heated films. This is a small effect probably dependent on evaporation rate, substrate material, and temperature.[9]* The low-energy peak is analogous to the first fundamental absorption of other alkali halides and exhibits a temperature dependence of 7×10^{-4} eV/°K. The second peak is similar to the first in all but one important aspect; its peak position shows no temperature dependence.

We note with interest the shoulder appearing at 6·3 eV (at 80°K) on the side of the third peak (which is off the diagram at this temperature). It cannot easily be resolved into a superposition of a few conventionally shaped peaks. (In other alkali iodides a similar shoulder is more conspicuous,[10]† the peak at 6 eV being absent.) This structure in CsI, which seems to shift toward lower $h\nu$ at higher temperatures, appears at roughly the same energy as the large external photo-emission. It is conceivable that this is the edge of a continuous absorption associated with band-to-band transitions. If such optical transitions take place from the top of the valence band to the bottom of the conduction band, we may tentatively estimate that the electron affinity of CsI is $\sim 0{\cdot}3$ eV or less.[11]‡

6. CONCLUSIONS

The salient result presented above is the abrupt rise in the photo-electric yield of CsI when $h\nu$ reaches 6 eV. We interpret this as emission from the valence band with an upper edge 6 eV below the vacuum level. Energy distributions of emitted electrons are consistent with this conclusion. Optical absorption measurements indicate faint structure at 6 eV possibly indicative of band-to-band transitions. We must then assume that the electron affinity is small, 0·3 eV or less.

Acknowledgements—We are indebted to P. D. JOHNSON for the use of his monochromator and to L. APKER for many stimulating conversations.

* HILSCH and POHL, *Z. Phys.* **57**, 145 (1929), found that thin layers of RbI pressed quartz between plates showed the first absorption peak shifted about 25 Å toward the red compared to as-evaporated films. These pressed layers probably represent a greater degree of crystal perfection. Our annealing process presumably improves the crystallinity of the evaporated film.

† This feature, which we suspect is an edge rather than a peak, is evident in the early work on RbI of HILSCH and POHL, reference 10.

‡ The situation is then similar to that for Cs_3Sb.

VOL.
1
1956/5

REFERENCES

1. APKER L. and TAFT E. *Phys. Rev.* **79,** 964 (1950), **81,** 698 (1951), **82,** 814 (1951).
2. SEITZ F. *Revs. Modern Phys.* **26,** 7 (1954). MOTT N. F. and GURNEY R. W. *Electronic Processes in Ionic Crystals* Chapter III. Oxford University Press, London (1940).
3. APKER L., TAFT E., and DICKEY J. *Phys. Rev.* **84,** 508 (1951). PHILIPP H. R. *Phys. Rev.* **94,** 777A (1954).
4. APKER L., TAFT E., and DICKEY J. *J. Opt. Soc. Amer.* **43,** 78 (1952).
5. APKER L. and TAFT E. *Phys. Rev.* **81,** 698 (1951).
6. APKER L., TAFT E., and DICKEY J. *Phys. Rev.* **76,** 270 (1949); **75,** 1181 (1949); **74,** 1462 (1948); **73,** 46 (1948).
7. JOHNSON P. D. *J. Opt. Soc. Amer.* **42,** 278 (1952).
8. HILSCH R. and POHL R. *Z. Phys.* **59,** 812 (1930). SCHNEIDER E. and O'BRYAN H. *Phys. Rev.* **51,** 293 (1937).
9. DELBECQ C. and YUSTER P. *J. Chem. Phys.* **22,** 921 (1954). FISCHER F. *Z. Phys.* **139,** 328 (1954).
10. HILSCH R. and POHL R. *Z. Phys.* **57,** 145 (1929) Fig. 3. MARTIENSSEN W. *Nachr. Akad. Wiss. Göttingen. Math.-Physik Kl.* 257 (1955). FESEFELDT H. *Z. Phys.* **64,** 626 (1930). FISCHER F. *Z. Phys.* **139,** 328 (1954).
11. BURTON J. A. *Phys. Rev.* **72,** 531A (1947). TAFT E. and APKER L. *J. Opt. Soc. Amer.* **43,** 81 (1953). SPICER W. E. *Bull. Amer. Phys. Soc.* **1,** 285 (1956).

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 164–174.

PARAMAGNETIC RESONANCE AND POLARIZATION OF NUCLEI IN METALS

I. M. LIFSHITZ, M. YA. AZBEL', AND V. I. GERASIMENKO

(*Received* 7 *August* 1956)

Physical-Technical Institute, Academy of Science of the Ukrainian S.S.R., Kharkoff

Abstract—A theory of paramagnetic resonance in metals is constructed, based on a simultaneous solution of MAXWELL'S equations and of the kinetic equation for the density operator. The polarization of the nuclei in the metal is determined. It is shown that this polarization always varies slowly with depth, diminishing as $\exp(-z/\delta_{\text{eff}})$, where $\delta_{\text{eff}} \sim v\sqrt{(t_0 T_f)} \sim 10^{-2}$ to 1 cm is the mean distance travelled by an electron between collisions in which its spin is reversed. It is found that in paramagnetic resonance there is a selective transparency of metallic films, and the transmitted wave is circularly polarized, if the constant magnetic field is normal to the surface. (The short communication about this theory is already published.[10])

VOL. 1 1956/5

1. INTRODUCTION

IT is known[1] that paramagnetic resonance of the conduction electrons occurs in a metal placed in a constant magnetic field H_0 and a variable electromagnetic field of frequency $\omega = \Omega_0 \equiv 2\mu H_0/\hbar$, where μ is the magnetic moment of the electron.

OVERHAUSER[2] has shown that this resonance is accompanied by a polarization of the nuclei in the metal, as if the nuclei had an effective magnetic moment μ_{eff} of

$$\mu_{\text{eff}} = \mu_{\text{nuc}} + \mu\alpha T_f/(1+\alpha T_f), \qquad (1)$$

where μ_{nuc} is the actual magnetic moment of the nucleus, T_f the time of free flight of an electron between collisions in which its spin is reversed, and $\alpha = (4\mu^2 H_1^2/\hbar^2)T_f$ is the probability per unit time that the electron's spin is reversed in the variable magnetic field $2H_1 \cos \omega t$. However, it is easily seen that formula (1) is applicable only for very thin metal samples, whose thickness d is small compared with the skin depth $\delta \sim 10^{-4}$ to 10^{-5} cm, for the resonance probability of spin reversal per unit time can be introduced only when the electron is in a nearly homogeneous field for a time considerably exceeding the field period. For a mass of metal this condition holds only if $\delta/v \gg 2\pi/\omega$, which corresponds to frequencies $\omega \gg (2\pi v/c)\sqrt{(2\pi\sigma/t_0)} \sim 10^{13}$ sec^{-1} and magnetic fields $H_0 = \hbar\omega/2\mu \gg 10^6$ oersted, which are practically unattainable at present. Hence the polarization of the nuclei by OVERHAUSER'S method occurs only in small particles, whose diameters are of the order of one micron.[3]

At the same time, we should expect that OVERHAUSER'S method would enable us to polarize nuclei in layers whose thickness is tens or hundreds of times the skin depth, for the polarization of the nuclei is given by the value of the magnetic moment M associated with the spins of the electrons. The polarization of an electron at any point is determined by the values of the magnetic field H_1 which it encounters in its path from the last collision with change of spin to the point considered. Hence the magnetic moment at this point will be related to the values of the magnetic field at all points within a distance of the order of δ_{eff}, the path travelled by an electron between two successive collisions with spin reversal. Since the time T_f between such collisions considerably exceeds the ordinary mean free flight time t_0 of the electron, the length δ_{eff} is $\delta_{\text{eff}} = v\sqrt{(\frac{1}{3}t_0 T_f)} \sim 10^{-2}$ to 1 cm, where v is the speed of the electron. Consequently, from quite low frequencies $\omega \gtrsim c^2/2\pi\delta_{\text{eff}}^2\sigma$, when $\delta_{\text{eff}} \gtrsim \delta$ (for $t_0 \sim 10^{-11}$ sec and $T_f \sim 10^{-6}$ sec, $\omega \gtrsim 10^3$ sec^{-1}), there is an "anomalous skin-effect" for the magnetic moment. The relation between the magnetic moment and

the variable magnetic field is an integral one, $\boldsymbol{M}(\boldsymbol{r}) = \int N(\boldsymbol{r}, \boldsymbol{r}')\boldsymbol{H}_1(\boldsymbol{r}')\,d\boldsymbol{r}'$, where the integration extends over a region of radius $\sim \delta_{\text{ffe}}$. This gives a very slow variation of the magnetic moment with depth, and the "skin depth" for the magnetic moment is δ_{eff}.

Two important physical consequences follow from this:

(1) The polarization of nuclei in a metal can be brought about in layers whose thickness is of the order of $\delta_{\text{eff}} \sim 10^{-2}$ to 1 cm. This enables us to obtain very thick polarized nuclear targets.*

(2) The slow decrease of the magnetic moment with depth means that, according to MAXWELL'S equations, $\boldsymbol{E}$ and $\boldsymbol{H}_1$ will have a small, slowly decreasing part. For a film of thickness $d \gg \delta$, this part will determine the coefficient of transmission of an electromagnetic wave through the film close to the resonance. Consequently, in paramagnetic resonance we have not only resonance absorption but also resonance transmission through the film, and the transmission coefficient may increase by many orders of magnitude. For example, at low temperatures the wave transmission coefficient through a film of thickness about 0·1 mm may be 10^{-11} to 10^{-16} at resonance, while elsewhere it is 10^{-40} to 10^{-50}. (We may mention that this phenomenon occurs at all temperatures.)

All the above topics are discussed in the present paper on the basis of a simultaneous solution of MAXWELL'S equations and the kinetic equation for the electron density operator. The polarization of the nuclei in the metal and the coefficient of transmission of an electromagnetic wave through a film are calculated.†

2. STATEMENT OF THE PROBLEM. THE COMPLETE SYSTEM OF EQUATIONS

To solve the problem, we have to find the electron density operator $\hat{f}$ (which is a c number with respect to co-ordinates and a q number with respect to spins). To do so, we write down the kinetic equation, which in the present case has the form‡

$$\frac{\partial \hat{f}}{\partial t} + \boldsymbol{v}\cdot\frac{\partial \hat{f}}{\partial \boldsymbol{r}} + \frac{\partial \hat{f}}{\partial \boldsymbol{p}}\cdot\left\{e\boldsymbol{E} + \frac{e}{c}\boldsymbol{v}\wedge\boldsymbol{H}\right\} + \frac{i}{\hbar}[\hat{\mathcal{H}},\hat{f}] + \left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = 0, \qquad (2)$$

$$\hat{\mathcal{H}} = \mu\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{H}, \qquad \boldsymbol{H} = \boldsymbol{H}_0 + \boldsymbol{H}_1(r,t),$$

where p, ϵ, $\boldsymbol{v} = \nabla_{\boldsymbol{p}}\epsilon(\boldsymbol{p})$ are the momentum, energy and velocity of the electron, $\boldsymbol{E}$ and $\boldsymbol{H}$ the electric and magnetic field intensities, $\hat{\boldsymbol{\sigma}}$ the spin operator,

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

the z axis is taken in the direction of the magnetic field H_0, and $(\partial f/\partial t)_{\text{coll}}$ is the collision integral; the force $\mu(\hat{\boldsymbol{\sigma}}\nabla)\boldsymbol{H}_1$ is omitted as negligibly small.

The operator $\hat{f}$ varies both because of the redistribution of the electrons in energy and momentum, and because of their redistribution between spins. Hence the collision integral is determined both by collisions accompanied by spin reversal (with relaxation time T_f) and by those in which only the energy and momentum change (with relaxation time t_0). As has been mentioned above, $T_f \gg t_0$, and so we can consider the two types of collision separately:

$$(\partial f/\partial t)_{\text{coll}} = (\partial f/\partial t)_{\text{co}} + (\partial f/\partial t)_{\text{sp}},$$

where $(\partial f/\partial t)_{\text{co}}$ is the collision integral for collisions without spin reversal, $(\partial f/\partial t)_{\text{sp}}$ is the collision integral for collisions with spin reversal. Without stating the exact form of these operators, we can write them down, in the usual way, in terms of the corresponding relaxation times t_0 and T_f. To do so, we must determine equilibrium operators corresponding to each type of collision separately.

Collisions with spin reversal in the absence of the variable field H_1, of course, give the equilibrium operator $\hat{f}_0$ in the constant magnetic field H_0:

* L. N. ROZENTSVEĬG and YA. M. FOGEL' have pointed out that the same method can be used to polarize nuclei of absorbed and adsorbed hydrogen.

† It may be remarked that the absorption of power from an electromagnetic wave incident on a metal in paramagnetic resonance, for a constant magnetic field perpendicular to the metal surface and a weak electromagnetic field (the resonance being unsaturated), has been determined by DYSON[4] from a study of electron diffusion. However, DYSON did not consider the polarization of nuclei and the selective transparency of films.

‡ After the present paper was finished, there appeared another[5] in which the same equation is derived.

$$\hat{f}_0 = \begin{bmatrix} f_0(\epsilon_0+\mu H_0) & 0 \\ 0 & f_0(\epsilon_0-\mu H_0) \end{bmatrix},$$

$$f_0(\epsilon) = (e^{(\epsilon-\zeta)/kT}+1)^{-1}. \tag{3}$$

Let us now determine the equilibrium operator $\hat{f}^\circ$ given by collisions without spin reversal. Since there is here no redistribution among spins, the probability density for a given electron spin projection is unchanged. Hence

$$\int \hat{f}\, d\tau_p = \int \hat{f}^\circ\, d\tau_p, \quad d\tau_p = dp_x\, dp_y\, dp_z. \tag{4}$$

For simplicity, we shall suppose that the energy change in collisions can be neglected. (This, of course, holds at low temperatures, when only the residual resistance is of importance and collisions can be regarded as elastic, and for weak magnetic fields, when $\mu H_0 \ll kT$.) Then, from (4)

$$\hat{f}^\circ = \int \hat{f}\, ds/v \Big/ \int ds/v \equiv \bar{\hat{f}}.$$

Thus equation (2) takes the form

$$\frac{\partial \hat{f}}{\partial t} + \boldsymbol{v}\cdot\frac{\partial \hat{f}}{\partial \boldsymbol{r}} + \frac{\partial \hat{f}}{\partial \boldsymbol{p}}\cdot\left\{e\boldsymbol{E}+\frac{e}{c}\boldsymbol{v}\wedge\boldsymbol{H}\right\} + \frac{i}{\hbar}[\mu\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{H}, \hat{f}] + \frac{\hat{f}-\bar{\hat{f}}}{t_0} + \frac{\hat{f}-\hat{f}_0}{T_f} = 0. \tag{5}$$

Let us now write down the boundary condition for the function $\hat{f}$ at the surface. The reflection of electrons from the surface can be described in a semiphenomenological manner[6] by a reflection coefficient q. Assuming that the spin of the electron is unchanged in collisions with the surface, we have

$$[\hat{f}]_{\boldsymbol{v}\cdot\boldsymbol{n}>0,\,\text{surface}} = (1-q)[\bar{\hat{f}}]_{\text{surface}} + q[\hat{f}]_{-\boldsymbol{v}\cdot\boldsymbol{n}<0,\,\text{surface}}, \tag{6}$$

where $\boldsymbol{n}$ is the inward normal to the surface.* At infinity, where the variable field H_1 vanishes, $\hat{f}$ reduces to $\hat{f}_0$. The boundary condition with respect to p is given by the requirement of a unique solution (see below).

* In the paper by SILIN[5] the condition corresponding to $q = 1$ is given. It seems that in practice we always have $q \simeq 0$. It may be mentioned that the conservation of the electron spin in a collision with the surface follows from (4) and (6).

Equation (5), with the boundary condition (6) and MAXWELL'S equations

$$\mathbf{curl}\, \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}, \quad \mathbf{curl}\, \boldsymbol{H}_1 = \frac{4\pi}{c}\boldsymbol{j}, \tag{7}$$

where

$$\boldsymbol{j} = \frac{e}{\hbar^3}\int \boldsymbol{v}\, \mathrm{tr}\hat{f}\, d\tau_p, \quad \boldsymbol{B} = \boldsymbol{H}_1 + 4\pi\boldsymbol{M},$$

$$\boldsymbol{M} = \frac{\mu}{\hbar^3}\int \mathrm{tr}(\hat{\boldsymbol{\sigma}}\hat{f})\, d\tau_p, \tag{8}$$

forms a complete system of equations for the problem considered.

It is convenient to expand the operators $\hat{f}$ and $\hat{f}_0$ in terms of the unit operator $\hat{I}$ and the spin operator $\hat{\boldsymbol{\sigma}}$:

$$\hat{f} = f_I\hat{I} + \boldsymbol{f}\cdot\hat{\boldsymbol{\sigma}}, \quad \hat{f}_0 = \hat{f}_{0I}\hat{I} + \boldsymbol{f}_0\cdot\hat{\boldsymbol{\sigma}}.$$

Then

$$\boldsymbol{j} = \frac{2e}{\hbar^3}\int \boldsymbol{v} f_I\, d\tau_p, \quad \boldsymbol{M} = \frac{2\mu}{\hbar^3}\int \boldsymbol{f}\, d\tau_p. \tag{9}$$

The function f_I and the relation between $\boldsymbol{j}$ and $\boldsymbol{E}$ given by (9) may be found in Reference 7. We shall be interested only in the function $\boldsymbol{f}$, which satisfies the equation

$$\frac{\partial \boldsymbol{f}}{\partial t} + v_\zeta\frac{\partial \boldsymbol{f}}{\partial \zeta} + \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{p}}\cdot\frac{e}{c}\boldsymbol{v}\wedge\boldsymbol{H} + \frac{\boldsymbol{f}-\boldsymbol{a}\chi_0}{T_f} + \frac{\boldsymbol{f}-\bar{\boldsymbol{f}}}{t_0} + \boldsymbol{f}\wedge\Omega = 0, \tag{10}$$

with the boundary condition

$$[\boldsymbol{f}]_{\boldsymbol{v}\cdot\boldsymbol{n}>0,\,\text{surface}} = (1-q)[\bar{\boldsymbol{f}}]_{\text{surface}} + q[\boldsymbol{f}]_{-\boldsymbol{v}\boldsymbol{n}<0,\,\text{surface}}.$$

Here

$$\boldsymbol{a} = \boldsymbol{H}_0/H_0 = (0,0,1); \quad \Omega = 2\mu\boldsymbol{H}/\hbar = \Omega_0+\Omega_1;$$

$$\chi_0 = \tfrac{1}{2}[f_0(\epsilon_0+\mu H_0) - f_0(\epsilon_0-\mu H_0)],$$

and ζ is the distance from the surface of the metal.†

† We shall use two systems of co-ordinates: ξ, y, ζ, a fixed system with ξ and y in the surface of the metal, and x, y, z, which is attached to the constant field H_0.

VOL. 1 1956/57

(We may notice that, in the general case of arbitrary temperatures and magnetic fields,

$$\left(\frac{\partial f}{\partial t}\right)_{\text{co}} = \frac{\bar{f}-f}{t_p} + \frac{\tilde{f}-\tilde{\chi}^\circ f/\chi^\circ}{t_\epsilon}, \quad \tilde{\phi} = \int \phi \, d\tau_p,$$

$$\chi^\circ = \tfrac{1}{2}[f_0(\epsilon_0+\Delta) - f_0(\epsilon_0-\Delta)],$$

where t_ϵ and t_p are the relaxation times for energy and momentum, and Δ is found from the equation $|\tilde{f}| = \tilde{\chi}^\circ$.

As the variables in momentum space we take the energy ϵ, the projected momentum p_z, and the dimensionless time of revolution of the electron in its orbit $\tau = (1/2\pi m_0)\partial S_\tau/\partial\epsilon$, where S_τ is the area of a sector in the intersection of the surface $\epsilon(p) = \epsilon$ and the plane p_z = constant.[8]
Then we have

$$\frac{\partial f}{\partial t} + v_\zeta \frac{\partial f}{\partial \zeta} + \frac{1}{T_0}\frac{\partial f}{\partial \tau} + \frac{f-a\chi_0}{T_f} + \frac{f-\bar{f}}{t_0} + f \wedge \Omega = 0. \tag{11}$$

Here $T_0 = m_0 c/eH_0$. For simplicity, we suppose that $\mu H_0 \ll kT$. Then, putting $f = -\mu H_0(\partial f_0/\partial\epsilon)(a-w)$, we obtain

$$M = \chi H_0(a - \bar{w}), \tag{12}$$

where χ is the static magnetic susceptibility of the metal, and w satisfies the equation

$$\frac{\partial w}{\partial t} + v_\zeta \frac{\partial w}{\partial \zeta} + \frac{1}{T_0}\frac{\partial w}{\partial \tau} + \frac{w}{t_0^*} + w \wedge \Omega = \frac{\bar{w}}{t_0} + a \wedge \Omega_1,$$

$$1/t_0^* = 1/t_0 + 1/T_f, \quad \bar{w} = \int w \, ds/v \Big/ \int ds/v, \tag{13}$$

with the boundary condition

$$[w]_{v,n>0,\text{surface}} = (1-q)[\bar{w}]_{\text{surface}} + q[w]_{-v,n<0,\text{surface}}. \tag{14}$$

At infinity, $(w)_{v,n<0} = 0$. The function w must be periodic in τ, with period $\theta = (1/2\pi m_0)\partial S/\partial\epsilon$ (cf. reference 8). Solving (13) and (14) with MAXWELL'S equations (7), we can determine the magnetic moment M and the electric and magnetic field intensities within the metal. This makes it possible to calculate the surface impedance of the metal, $Z = (4\pi/c)E(0)/H_1(0)$, the coefficient of transmission of an electromagnetic wave through the film, and the polarization of the nuclei[2]

$$P = \frac{1}{I}\{(I+\tfrac{1}{2})\coth(I+\tfrac{1}{2})\bar{w}_z - \tfrac{1}{2}\coth\tfrac{1}{2}\bar{w}_z\}, \tag{15}$$

where I is the spin of the nucleus.

3. REDUCTION OF THE EQUATIONS TO CANONICAL FORM

To solve equations (7) and (13) we put

$$w = -w_y - iw_x, \quad \Omega_1 = \Omega_{1x} + i\Omega_{1y}, \quad M = M_x + iM_y.$$

Then, if $H_{1z} \ll H_0$ ($H_{1z} \equiv 0$ if H_0 is perpendicular to the metal surface), we have

$$\left.\begin{aligned} \left(\frac{\partial}{\partial t} + \hat{D} - i\Omega_0\right)w &= \frac{\bar{w}}{t_0} + \Omega_1(1-w_z), \\ \left(\frac{\partial}{\partial t} + \hat{D}\right)w_z &= \frac{\bar{w}_z}{t_0} + \operatorname{re}(\Omega_1^* w), \\ \hat{D} &\equiv v_\zeta\frac{\partial}{\partial\zeta} + \frac{1}{T_0}\frac{\partial}{\partial\tau} + \frac{1}{t_0^*}, \end{aligned}\right\} \tag{16}$$

$$M_z = \chi H_0(1-\bar{w}_z), \quad M = i\chi H_0 \bar{w}. \tag{17}$$

We consider only the component of Ω_1 which gives the resonance, i.e. $\Omega_1 = \Omega_1 e^{i\omega t}$ (denoting the amplitude by the same letter as the function itself). Then the solution is $w = we^{i\omega t}$ and w_z independent of time. The equations determining $w(\zeta)$ and $w_z(\zeta)$ are

$$\left.\begin{aligned} \{\hat{D} + i(\omega - \Omega_0)\}w &= \bar{w}/t_0 + \Omega_1(1-w_z), \\ \hat{D}w_z &= \bar{w}_z/t_0 + \operatorname{re}(\Omega_1^* w). \end{aligned}\right\} \tag{18}$$

Thus the problem reduces to the solution of the system of equations (18) and MAXWELL'S equations (7).

It is evident that this system of equations is in general nonlinear, because of the nonlinear relation between the magnetic moment M and the field $H_1 = \hbar\Omega_1/2\mu$ (this relation is given by the equations 18).

The most interesting case is that of a fairly strong field H_1, when the electron gas is almost completely depolarized near the surface of the metal ($w_z \approx 1$), while at fairly large depths the depolarization is very small ($w_z \ll 1$), so that no ordinary linearization is possible.

To solve the equations, we use a fact mentioned

in Section 1, that the magnetic moment, and therefore w and w_z, decreases considerably more slowly than Ω_1 at moderate distances from the surface. (At distances much greater than δ_{eff}, Ω_1 itself varies as slowly as M.)

We shall first explain the mathematical reason for the slow decrease of M. To do so we consider an "ideal resonance", where $w = \Omega_0$ and $T_f = \infty$. Then equations (18), in the absence of the variable magnetic field, i.e. $\Omega_1 = 0$, are the same for w and w_z:

$$\left.\begin{aligned} &\left(v_\zeta\frac{\partial}{\partial\zeta}+\frac{1}{T_0}\frac{\partial}{\partial\tau}+\frac{1}{t_0}\right)w=\frac{\bar{w}}{t_0},\\ &[w]_{v,\,n>0,\text{surface}}=(1-q)[\bar{w}]_{\text{surface}}+\\ &\qquad\qquad+q[w]_{-v,\,n<0,\text{surface}}. \end{aligned}\right\} \tag{19}$$

This homogeneous equation has a nontrivial solution in the form of a function of ϵ only: $w = \bar{w}$ = constant $\phi(\epsilon)$ (i.e. a constant with respect to τ and ζ). For $\omega = \Omega_0$, therefore, $t_0/t_0^* = 1$ is an eigenvalue of equation (19), and the corresponding eigenfunction is independent of the coordinates. Since in equations (18) $t_0/t_0^* = 1 + t_0/T_f \approx 1$, the solution is, firstly, considerably greater than Ω_1 (since for $t_0/T_f = 0$ the equations have no solution except those diverging as $(t_0/T_f)^{-\frac{1}{2}}$), and therefore nearly equal to $\bar{w}$ and $\bar{w}_z$; secondly, it is nearly equal to the eigenfunction, i.e. it varies only slowly with distance; thirdly, it is affected by the behaviour of Ω_1 only at small distances (since the equation has a smoothly varying solution even if Ω_1 is a delta function of the coordinates, $\Omega_1 = \delta(\zeta)$).

Accordingly we make the following verifiable assumptions:

(1) The quantities w and w_z differ only slightly from their mean values: $|w-\bar{w}| \ll \bar{w}$, $|w_z-\bar{w}_z| \ll \bar{w}_z$. In equations (18) we can therefore put $\Omega_1(1-w_z) \approx \Omega_1(1-\bar{w}_z)$ and $\Omega_1^* w \approx \Omega_1^* \bar{w}$.

(2) The values of w and w_z vary greatly with the co-ordinate ζ only at distances δ_{eff} which considerably exceed the skin depth δ. At distances small compared with δ_{eff}, therefore, $\Omega_1(1-\bar{w}_z) \approx \Omega_1(1-\bar{w}_z(0))$ and $\Omega_1^*\bar{w} \approx \Omega_1^*\bar{w}(0)$.

(3) The behaviour of the inhomogeneity in equations (18) at large distances is unimportant.†

† This assumption is essential, because at large distances Ω_1 is determined principally by a small but slowly decreasing part related to M; in this region $\Omega_1 w \ll \Omega_1 w(0)$.

Thus equations (18) near the resonance can be written

$$\left.\begin{aligned} \{\hat{D}+i(\omega-\Omega_0)\}w &= \bar{w}/t_0+\Omega_1[1-\bar{w}_z(0)],\\ \hat{D}w_z &= \bar{w}_z/t_0+\text{re}[\Omega_1^*\bar{w}(0)]. \end{aligned}\right\} \tag{20}$$

Putting

$$w=(1-\bar{w}_z(0))u, \quad w_z=\text{re}\,(\bar{w}(0)u_z), \tag{21}$$

equations (20) become

$$\left.\begin{aligned} \{\hat{D}+i(\omega-\Omega_0)\}u &= \bar{u}/t_0+\Omega_1,\\ \hat{D}u_z &= \bar{u}_z/t_0+\Omega_1. \end{aligned}\right\} \tag{22}$$

Averaging, we easily find from (21)

$$\bar{w}_z(0)=\frac{\text{re}\{\bar{u}(0)\bar{u}_z(0)\}}{1+\text{re}\{\bar{u}(0)\bar{u}_z(0)\}},\ \bar{w}(0)=\frac{\bar{u}(0)}{1+\text{re}\{\bar{u}(0)u_z(0)\}},$$

and therefore

$$w_z(\zeta)=\frac{\text{re}\{u(0)u_z(\zeta)\}}{1+\text{re}\{\bar{u}(0)\bar{u}_z(0)\}},\ w(\zeta)=\frac{u(\zeta)}{1+\text{re}\{u(0)\bar{u}_z(0)\}}. \tag{23}$$

The magnetization M_z and M and the polarization P of the nuclei are determined as before by formulae (17) and (15).

It is seen from equations (22) that $u_z(\zeta) = (u^*(\zeta))_{\omega=\Omega_0}$, and therefore it suffices to solve only the equation‡

$$v_\zeta\frac{\partial u}{\partial\zeta}+\frac{1}{T_0}\frac{\partial u}{\partial\tau}+\frac{u}{t_0^*}+i(\omega-\Omega_0)u=\frac{\bar{u}}{t_0}+\Omega_1, \tag{24}$$

with the boundary condition

$$\left.\begin{aligned} [u]_{v,\,n>0,\text{surface}} &= (1-q)[u]_{\text{surface}}+\\ &\quad+q[u]_{-v,\,n<0,\text{surface}}. \end{aligned}\right\} \tag{25}$$

At infinity $(u)_{v,n<0} = 0$. The boundary condition with respect to τ is given by the requirement that u is periodic with period $\theta = (1/2\pi m_0)\partial S/\partial\epsilon$.

4. THE SOLUTION OF THE EQUATIONS FOR A MASS OF METAL

(a) *A particular case*

Let us consider what is mathematically the simplest case, that of a quadratic dispersion

‡ We may note that for this reason, at resonance,

$$w_z(\xi)=\frac{\text{re}\{\bar{u}(0)u^*(\zeta)\}}{1+|\bar{u}(0)|^2},\ w(\xi)=\frac{u(\xi)}{1+|\bar{u}(0)|^2}.$$

VOL. 1 1956/5

$\epsilon = p^2/2m^*$ (m^* being the effective mass), a field H_0 perpendicular to the metal surface (the coordinate systems $\xi y \zeta$ and xyz being here the same), and specular reflection of electrons from the surface, i.e. $q = 1$. (This case is, of course, of purely academic interest.) Then equation (24) takes the form

$$v_z \frac{\partial u}{\partial z} + \frac{u}{t} = \frac{\tilde{u}}{t_0} + \Omega_1, \quad \frac{1}{t} = \frac{1}{t_0} + \frac{1}{T_f} + i(\omega - \Omega_0). \tag{26}$$

The boundary conditions are

$$[u]_{v_z>0, z=0} = [u]_{-v_z<0, z=0}, \quad [u]_{v_z<0, z=\infty} = 0.$$

Solving this equation and averaging it over the Fermi surface, we obtain an integral equation for $\tilde{u}$:

$$\tilde{u}(z) = \int_{-\infty}^{\infty} M(|z-\zeta|) \left\{ \frac{t}{t_0} \tilde{u}(\zeta) + t\Omega_1(\zeta) \right\} d\zeta, \tag{27}$$

where

$$M(|z-\zeta|) = \frac{1}{2l} \int_1^{\infty} e^{-|z-\zeta|x/l} \frac{dx}{x},$$

$$l = v_0 t, \ \Omega_1(-\zeta) = \Omega_1(\zeta).$$

We may notice that, for $\Omega_1 = 0$ and $t/t_0 = 1$, $\tilde{u}$ = constant satisfies (27).

From (27) we find by a Fourier transformation

$$\tilde{u}(z) = \frac{1}{\pi} \int_0^{\infty} \frac{M_k t_0 \Omega_{1k} \cos kz \, dk}{t_0/t - M_k}, \tag{28}$$

where

$$M_k = \frac{\tan^{-1} kl}{kl}, \quad \Omega_{1k} = 2\int_0^{\infty} \Omega_1(z) \cos kz \, dz.$$

Close to the resonance ($|\omega - \Omega_0| \sim 1/T_f$), only small values of k (such that $kl \ll 1$) are important in the integral (28). For such kl, $M_k \approx 1 - \frac{1}{3}k^2l^2$ and the denominator of the integrand is $t_0/T_f + it_0(\omega - \Omega_0) + \frac{1}{3}k^2l^2$, while for $kl \gg 1$ it is $1 + t_0/T_f + it_0(\omega - \Omega_0) + \pi/2kl$. It is easy to see that the main contribution to the integral is from $kl \lesssim (t_0/T_f)^{\frac{1}{2}} \ll 1$. Then, close to the resonance,

$$\tilde{u}(z) = \frac{1}{\pi} \int_0^{\infty} \frac{t_0 \Omega_{10} \cos kz \, dk}{t_0/T_f + it_0(\omega - \Omega_0) + \frac{1}{3}k^2l^2}.$$

As we shall show below, the solution of MAXWELL's equations for $\boldsymbol{M} = 0$ can be used for Ω_1. Hence

$$\Omega_{1k} = 2\int_0^{\infty} \Omega_1(z) \cos kz \, dz$$

$$= \frac{4\mu c}{\hbar\omega} \{E(0) - k\int_0^{\infty} E(z) \sin kz \, dz\}$$

and

$$\Omega_{10} = 4\mu c E(0)/\hbar\omega.$$

Here $E(0)$ is the field at the surface just inside the metal.

Thus, close to the resonance,

$$\tilde{u}(z) = \frac{t_0 \Omega_{10}}{2l} \sqrt{\frac{3}{\kappa}} e^{-\sqrt{(3\kappa)}z/l}, \quad \kappa = \frac{t_0}{T_f} + it_0(\omega - \Omega_0).$$

It is seen from this formula that the width of the resonance line is determined only by the value of T_f : $|\omega - \Omega_o| \sim 1/T_f$, as was first shown by DYSON.[4]

At resonance,

$$\tilde{u}(z) = \frac{cE(0)}{H_0 v_0} \sqrt{\frac{3T_f}{t_0}} e^{-z/\delta_{\text{eff}}}, \ \delta_{\text{eff}} = v_0 \sqrt{(\tfrac{1}{3} t_0 T_f)}. \tag{29}$$

For the magnetization at resonance we find, substituting the value of ($\tilde{u}z$) in formulae (23),

$$M_z = \chi H_0 \left\{ 1 - \frac{|\alpha|^2}{1+|\alpha|^2} e^{-z/\delta_{\text{eff}}} \right\},$$

$$M = M_x + iM_y = i\chi H_0 \frac{\alpha}{1+|\alpha|^2} e^{-z/\delta_{\text{eff}}},$$

where

$$\alpha = \frac{c\{E_x(0) + iE_y(0)\}}{H_0 v_0} \sqrt{\frac{3T_f}{t_0}}.$$

For a sufficiently weak field H_1, when there is no saturation, i.e. $|\alpha| \ll 1$, the formulae for M_x and M_y become, apart from the factor $e^{-z/\delta_{\text{eff}}}$, formula (31) of reference 4, where we must put $\nu = \omega$.

Thus we find that $\tilde{u}(z)$ in fact decreases slowly. Moreover, u and $\tilde{u} \gg \Omega_1 t$, while $u - \tilde{u} \sim \Omega_1 t_o$, i.e. $u - \tilde{u} \ll \tilde{u}$.

We shall now show that the particular behaviour of $\Omega_1(z)$ at great depths is immaterial (i.e. we can use the solution of MAXWELL'S equations for $\boldsymbol{M} = 0$). In our formulae $\Omega_1(z)$ enters as $\int_0^{\infty} \Omega_1(z) \, dz$.

At distances greater than δ we have only the magnetic field due to the magnetization $\boldsymbol{M}$. Hence we can qualitatively represent $\Omega_1(z)$ as the sum of two parts: a large rapidly decreasing part and a small slowly decreasing part, i.e.

$$\Omega_1(z) \sim \frac{2\mu}{h}\left\{H_1 e^{-z/\delta} - 4\pi\chi H_0 \frac{i\alpha}{1+|\alpha|^2} e^{-z/\delta\sqrt{1}}\right\}.$$

Substituting this expression in the integral $\int_0^\infty \Omega_1(z)\,dz$, we easily find the condition for M to be negligible in MAXWELL's equations:

$$4\pi\chi\omega T_f \ll 1,$$

which in practice is always fulfilled. Thus the assumptions made in the preceding section are justified.

(b) *The general case*

We shall now find the quantities of interest for any dispersion law $\epsilon(p)$ and any magnitude and direction of the constant field $\boldsymbol{H}_0$. For simplicity, we shall consider only the case of resonance. We suppose the reflection of electrons from the surface to be diffuse ($q = 0$), as is always the case in practice.

It has been shown that the problem reduces to finding a solution, periodic in τ, of the equation

$$v_\zeta \frac{\partial u}{\partial \zeta} + \frac{1}{T_0}\frac{\partial u}{\partial \tau} + \frac{u}{t_0^*} = \frac{\tilde{u}}{t_0} + \Omega_1, \tag{30}$$

with the boundary conditions

$$[u(0)]_{v,n>0} = \tilde{u}(0), \quad [u(\infty)]_{v,n<0} = 0. \tag{31}$$

We put

$$\xi = \zeta/r_0, \quad \gamma = T_0/t_0, \quad \gamma^* = T_0/t_0^*,$$
$$v_\zeta/v_0 = V_\zeta, \quad r_0 = v_0 T_0,$$

where v_0 is the characteristic speed at the limiting Fermi surface. Then equation (30) becomes

$$\frac{\partial u}{\partial \xi} + \hat{L}u = \frac{\gamma}{V_\zeta}(\tilde{u} + t_0\Omega_1) \equiv \frac{\gamma}{V_\zeta}\psi,$$
$$\hat{L} \equiv \frac{1}{V_\zeta}\left(\frac{\partial}{\partial\tau} + \gamma^*\right). \tag{32}$$

To solve this equation we use the method developed in (9). We replace $\boldsymbol{p}$ in (32) by $-\boldsymbol{p}$; we then obtain for the function $u(-\boldsymbol{v})$ the equation

$$\frac{\partial u(-\boldsymbol{v})}{\partial \xi} - \hat{L}u(-\boldsymbol{v}) = -\frac{\gamma}{V_\zeta}\psi. \tag{33}$$

(We here use the central symmetry of the Fermi surface: $\epsilon(-\boldsymbol{p}) = \epsilon(\boldsymbol{p})$, $\boldsymbol{v}(-\boldsymbol{p}) = -\boldsymbol{v}(\boldsymbol{p})$,

$$\left[\frac{1}{\tau_0}\frac{\partial}{\partial\tau}\right]_{\boldsymbol{p}} \equiv \boldsymbol{v}\wedge\frac{\partial}{\partial\boldsymbol{p}} = \left[\frac{1}{T_0}\frac{\partial}{\partial\tau}\right]_{-\boldsymbol{p}}.)$$

Operating on equation (32) with $\partial/\partial\xi - \hat{L}$, and on (33) with $\partial/\partial\xi + \hat{L}$, and subtracting, we obtain an equation for the function $f = \frac{1}{2}(u(\boldsymbol{v}) + u(-\boldsymbol{v}))$:

$$\left(\frac{\partial^2}{\partial\xi^2} - \hat{L}^2\right) f = -2\gamma\psi\hat{L}\frac{1}{V_\zeta}, \tag{34}$$

where we have used the fact that $\bar{f}(\xi) = \tilde{u}(\xi)$.

In this equation we extend the functions $f(\xi)$ and $\Omega_1(\xi)$ to $\xi < 0$ so as to be even: $f(-\xi) = f(\xi)$, $\Omega_1(-\xi) = \Omega_1(\xi)$, and take the Fourier transform with

$$\phi_k = \int_{-\infty}^{\infty} f(\xi) e^{ik\xi}\,d\xi, \quad \Omega_{1k} = \int_{-\infty}^{\infty} \Omega_1(\xi) e^{ik\xi}\,d\xi$$

(clearly $\bar{\phi}_k = \tilde{u}_k$):

$$(\hat{L}^2 + k^2)\phi_k = \gamma\psi_k\hat{L}\frac{1}{V_\zeta} - 2f'(0). \tag{35}$$

From equations (32) and (33) we find

$$f'(0) = -\tfrac{1}{2}\hat{L}\{u(\boldsymbol{v}) - u(-\boldsymbol{v})\}.$$

We now use the boundary condition (31). Since, for $V_\zeta > 0$, $u = \tilde{u} = \bar{f}$ at the metal surface independently of the other components of velocity, it is easy to see that

$$u(\boldsymbol{v}) - u(-\boldsymbol{v}) = \operatorname{sign} V_\zeta \cdot \{\bar{f}(0) - f(0)\},$$

and

$$f'(0) = -\operatorname{sign} V_\zeta \cdot \hat{L}\{\bar{f}(0) - f(0)\}.$$

Hence equation (35) becomes

$$(\hat{L}^2 + k^2)\phi_k = 2\hat{L}\frac{1}{V_\zeta}\{\tfrac{1}{2}\gamma\psi_k + |V_\zeta|[\bar{f}(0) - f(0)]\},$$

and

$$\phi_k = \{(\hat{L}+ik)^{-1}+(\hat{L}-ik)^{-1}\} \cdot \times$$
$$\times \frac{1}{V_\zeta}\{\tfrac{1}{2}\gamma\psi_k + |V_\zeta|[\bar{f}(0)-\bar{f}(0)]\} \equiv g_+ + g_-.$$

The right-hand side of this expression is found by determining the periodic solution of the linear equation

$$\left(\frac{\partial}{\partial\tau}+\gamma^* \pm ikV_\zeta\right) g_\pm = \tfrac{1}{2}\gamma\psi_k + |V_\zeta|[\bar{f}(0)-f(0)],$$

whence

$$g_\pm = \frac{1}{e^{\gamma^*\theta \pm ik\tilde{V}_\zeta}-1}\int_\tau^{\tau+\theta} e^{\gamma^*(\tau'-\tau)\pm ik\int_\tau^{\tau'} V_\zeta\, d\tau''} \times$$
$$\times\{\tfrac{1}{2}\gamma\psi_k + |V_\zeta|[\bar{f}(0)-f(0)]\}\, d\tau', \quad \tilde{V}_\zeta = \int_0^\theta V_\zeta\, d\tau.$$

Then, substituting the expression for ψ_k, we have

$$\phi_k(\tau) = M(k,\tau)u_k + M(k,\tau)t_0\Omega_{1k} + \int_\tau^{\tau+\theta} N(k,\tau,\tau')[\bar{f}(0)-f(0)]\, d\tau', \qquad (36)$$

where

$$M(k,\tau) = \mathrm{re}\frac{\gamma}{e^{\gamma^*\theta+ik\tilde{V}_\zeta}-1}\int_\tau^{\tau+\theta} e^{\gamma^*(\tau'-\tau)+ik\int_\tau^{\tau'} V_\zeta d\tau''}\, d\tau',$$

$$N(k,\tau,\tau') = 2\,\mathrm{re}\frac{|V_\zeta|}{e^{\gamma^*\theta+ik\tilde{V}_\zeta}-1}\, e^{\gamma^*(\tau'-\tau)+ik\int_\tau^{\tau'} V_\zeta d\tau''}$$

We notice that, although $\bar{f}-f \ll \bar{f}$, $\bar{f}-f \sim \Omega_1 t_0$. Hence this difference cannot be neglected.

For convenience, we put

$$S_k(\tau) = \bar{\phi}_k(\tau) - \phi_k(\tau),$$

$$S(\tau) = \bar{f}(0,\tau) - f(0,\tau) = \lim_{\xi\to 0}\frac{1}{\pi}\int_0^\infty S_k(\tau)\cos k\xi\, dk$$
$$= \frac{1}{\pi}\int_0^\infty S_k(\tau)\, dk.$$

Then equation (36) becomes

$$S_k(\tau) = [1-M(k,\tau)]u_k - M(k,\tau)t_0\Omega_{1k} - \int_\tau^{\tau+\theta} N(k,\tau,\tau')S(\tau')\, d\tau'. \qquad (37)$$

Averaging this equation over the Fermi surface, we find the formation $\bar{u}_k$ (the required solution $\tilde{u}(\xi) = \frac{1}{\pi}\int_0^\infty u_k \cos k\xi\, dk$):

$$\bar{u}_k = \frac{\overline{M(k,\tau)}\cdot t_0\Omega_{1k}}{1-\overline{M(k,\tau)}} + \frac{1}{1-\overline{M(k,\tau)}}\overline{\int_\tau^{\tau+\theta} N(k,\tau,\tau')S(\tau')\, d\tau'}. \qquad (38)$$

To determine $S(\tau)$, we substitute $\bar{u}_k$ in (37) and integrate over k from 0 to ∞, obtaining the integral equation

$$S(\tau) = \frac{1}{\pi}\int_0^\infty \frac{\overline{M(k,\tau)}-M(k,\tau)}{1-\overline{M(k,\tau)}} \times$$
$$\times t_0\Omega_{1k}\, dk - \frac{1}{\pi}\int_0^\infty dk \int_\tau^{\tau+\theta} N(k,\tau,\tau')S(\tau')\, d\tau' +$$
$$+\frac{1}{\pi}\int_0^\infty \frac{1-M(k,\tau)}{1-\overline{M(k,\tau)}}\, dk \overline{\int_\tau^{\tau+\theta} N(k,\tau,\tau')S(\tau')\, d\tau'}.$$

Similarly to the procedure used in reference 9, we can solve this equation by successive approximations and show that $S(\tau) \sim t_0\Omega_1 k$ has no singularities for any τ.

Returning to (38), we notice that, for sufficiently small k (such that $kl \ll 1$),

$$\overline{M(k,\tau)} = t_0^*/t_0 - l^2k^2,$$

where

$$l^2 = \frac{\gamma}{\theta\int dp_z}\int \frac{dp_z}{(e^{\gamma^*\theta}-1)^3} \times$$
$$\times \int_0^\theta d\tau \int_0^\theta e^{\gamma^*\tau'}\phi(\tau,\tau')\, d\tau',$$

$$\phi(\tau,\tau') = e^{\gamma^*\theta}(e^{\gamma^*\theta}+1)\tilde{V}_\zeta^2 - 2\int_\tau^{\tau+\tau'} V_\zeta\, d\tau'' \cdot e^{\gamma^*\theta}(e^{\gamma^*\theta}-1)\tilde{V}_\zeta +$$
$$+\left(\int_\tau^{\tau+\tau'} V_\zeta d\tau''\right)^2 (e^{\gamma^*\theta}-1)^2.$$

It is easily seen that, for weak fields $H_0(\gamma \gg 1)$,

$l \sim 1/\gamma$. For strong fields $H_0(\gamma \ll 1)$, two cases are possible:

(1) If the field H_0 makes with the surface an angle $\gg \gamma$ (when $\tilde{V}_\zeta$ is noticeably different from zero), then $\varphi(\tau, \tau') \simeq 2\tilde{V}_\zeta^2$, and $l \sim 1/\gamma$ as before.

(2) If the field H_0 makes with the surface an angle $\lesssim \gamma$ (when $\tilde{V}_\zeta \simeq 0$), then $l \sim 1$.

Similarly to (a), we have

$$\tilde{u}(\xi) = \frac{1}{\pi}\{t_0\Omega_{10} + \int_\tau^{\tau+\theta} N(0,\tau,\tau')S(\tau')\,d\tau'\} \cdot \int_0^\infty \frac{\cos k\xi\, dk}{t_0/T_f + l^2k^2}.$$

From MAXWELL's equations for $\boldsymbol{M} = 0$, $\Omega_{10} = (4\mu c/\hbar\omega)E(0)$. Taking into account the fact that $S(\tau) \sim r_0\Omega_{10}$, we find

$$\tilde{u}(\xi) = A\frac{ct_0E(0)}{H_0r_0l}\sqrt{\frac{T_f}{t_0}}e^{-\zeta/\delta_{\text{eff}}}, \qquad (39)$$

where

$$\delta_{\text{eff}} = lr_0\sqrt{\frac{T_f}{t_0}}, \quad A \sim 1, \quad E(0) = \frac{cZ}{4\pi}H_1(0),$$

and Z is the surface impedance without allowing for the magnetization $\boldsymbol{M}$, which has been found in reference 7.

The exact value of l depends on the dispersion law and on the direction of the field $\mathbf{H}_0$. For instance, for a quadratic dispersion and a constant field perpendicular to the metal surface, $l = 1/\gamma\sqrt{3}$ and we recover formula (29).

Using formulae (23) and (15), we find for the polarization of the nuclei in a mass of metal, at resonance,

$$P = \frac{1}{I}\left\{(I+\tfrac{1}{2})\coth(I+\tfrac{1}{2})\frac{|\alpha|^2e^{-\zeta/\delta_{\text{eff}}}}{1+|\alpha|^2}\frac{\mu H_0}{kT} - \right.$$
$$\left. -\frac{1}{2}\coth\frac{1}{2}\frac{|\alpha|^2e^{-\zeta/\delta_{\text{eff}}}}{1+|\alpha|^2}\frac{\mu H_0}{kT}\right\},$$

where

$$\alpha = A\frac{ct_0E(0)}{H_0r_0l}\sqrt{\frac{T_f}{t_0}}.$$

Thus, in a mass of metal, the nuclei can be polarized at a depth $\delta_{\text{eff}} \sim v_0T_0\sqrt{(T_f/t_0)}$ for a strong field H_0 parallel to the surface, and $\delta_{\text{eff}} \sim v_0t_0\sqrt{(T_f/t_0)}$ in all other cases.

5. THE SELECTIVE TRANSPARENCY OF THE FILM

In order to find the coefficient of transmission of an electromagnetic wave through a film of thickness d, it is necessary to solve equation (24) with the boundary conditions (25). For simplicity, we consider the case of a quadratic dispersion law $\epsilon = p^2/2m^*$, a field H_0 perpendicular to the surface, and specular reflection of the electrons from the surface (since, as we have seen, the actual nature of the dispersion law and the boundary conditions have no qualitative effect on the results). Then equation (24) becomes

$$v_z\frac{\partial u}{\partial z} + \frac{u}{t} = \frac{\tilde{u}}{t_0} + \Omega_1, \quad \frac{1}{t} = \frac{1}{t_0^*} + (\omega - \Omega_0). \qquad (40)$$

The boundary conditions are

$$[u]_{v_z>0,z=0} = [u]_{-v_z<0,z=0},$$
$$[u]_{v_z<0,z=d} = [u]_{-v_z>0,z=d}.$$

Having found $u(z)$ and averaged it over the Fermi surface, we find for $\tilde{u}(z)$ the integral equation

$$\tilde{u}(z') = \frac{1}{2v_0}\int_{-k}^{k} M(|z'-\zeta|)\{\tilde{u}(\zeta) + t_0\Omega_1(\zeta)\}\,d\zeta, \qquad (41)$$

where

$$M(\nu) = \frac{t}{t_0}\int_1^\infty \frac{\cosh(k-|\nu|)x}{\sinh kx}\frac{dx}{x}, \; z' = z/v_0t, \; k = d/v_0t.$$

The function $M(\nu)$ is even and periodic in the range $(-2k, 2k)$, the period being $2k$. The equation (41) can therefore be solved by expanding all functions in a Fourier series of cosines with period π/k. The result is

$$\tilde{u}(z') = \tfrac{1}{2}u_0 + \sum_{n=1}^\infty u_n\cos(\pi nz'/k), \qquad (42)$$

where

$$u_n = \frac{t_0\Omega_{1n}kM_n}{1-kM_n}, \quad \Omega_{1n} = \frac{2}{k}\int_0^k \Omega_1(z)\cos\frac{\pi nz}{k}dz,$$

$$M_n = \int_0^k M(z)\cos\frac{\pi nz}{k}dz = \frac{1}{k}\frac{t}{t_0}\frac{\tan^{-1}(\pi n/k)}{\pi n/k}.$$

At resonance, the important contribution to (42)

VOL. 1 1956/57

is from u_n with small n (such that $\pi n/k \ll 1$), since here

$$kM_n \simeq \frac{t_0^*}{t_0}\left\{1-\tfrac{1}{3}\left(\frac{\pi n}{k}\right)^2\right\}.$$

Then

$$u_n \simeq \frac{t_0\Omega_{1n}}{t_0/T_f+\frac{1}{3}(\pi n/k)^2}.$$

We see from this that only those n are important for which

$(\pi n/k)^2 \lesssim t_0/T_f \ll 1$, i.e. $n \lesssim (k/\pi)\sqrt{(t_0/T_f} = d/\pi\delta_{\text{eff}}$, where $\delta_{\text{eff}} = v_0\sqrt{(\frac{1}{3}t_0T_f)}$. Finally, since the important $\Omega_1(z)$ are those for which $z \lesssim \delta/v_0t_0$,

$$\Omega_{1n} \simeq \frac{2}{k}\int_0^k \Omega_1(z)\cos\sqrt{(t_0/T_f)}z\,dz$$

$$\simeq \frac{2}{k}\int_0^k \Omega_1(z)\,dz \equiv \Omega_{10}.$$

Thus

$$u_n \simeq \frac{t_0\Omega_{10}}{t_0/T_f+\frac{1}{3}(\pi n/k)^2}$$

and

$$\bar{u}(z) = \tfrac{1}{2}\Omega_{10}T_f\frac{d}{\delta_{\text{eff}}}\cdot\frac{\cosh[(d-z)/\delta_{\text{eff}}]}{\sinh(d/\delta_{\text{eff}})}. \qquad (43)$$

From MAXWELL'S equations, neglecting the magnetization $\boldsymbol{M}$, we have

$$\Omega_{10} = \frac{4\mu c}{\hbar\omega}\,\frac{E(0)-E(d)}{d}.$$

We may notice that formula (43) becomes (29) when $d \to \infty$.

It is very interesting to consider the selective transparency of films for which $\delta \ll d \ll \delta_{\text{eff}}$. Here (at resonance)

$$\bar{u}(z) \simeq cT_fE(0)/H_0d.$$

For the magnetization we have

$$M = i\chi H_0\frac{\bar{u}}{1+|\bar{u}|^2}e^{i\omega t} \equiv M_0e^{i\omega t}.$$

The transmission coefficient K is given by

$$K = |H_{1,\text{transmitted}}/H_{1,\text{incident}}|^2.$$

From MAXWELL'S equations (7) for

$$H_1 = H_{1x}+iH_{1y},\ E = E_x+iE_y,$$
$$j = j_x+ij_y,\ M = M_x+iM_y,$$

we obtain the equations

$$\frac{\partial E}{\partial z} = -\frac{\omega}{c}(H_1+4\pi M_0),\quad \frac{\partial H_1}{dz} = -\frac{4\pi i}{c}j.$$

Since, in a film of thickness $d \ll \delta_{\text{eff}}$, M is almost independent of z, we have, replacing H_1 by $H_{\|}-4\pi M_0$,

$$\frac{\partial E}{\partial z} = -\frac{\omega}{c}H_{\|},\quad \frac{\partial H_{\|}}{\partial z} = -\frac{4\pi i}{c}j,$$

which have rapidly diminishing solutions (with relaxation depth δ) for $H_{\|}$ and E.

Consequently, at depths considerably exceeding δ, there is only a homogeneous magnetic field $-4\pi M_0e^{i\omega t}$ (it is easily seen that this field is circularly polarized).

Thus, for films whose thickness satisfies the inequality $\delta \ll d \ll \delta_{\text{eff}}$,

$$K \sim \left|\frac{\chi T_fc^3Z^2}{2\pi d\left(1+\left|\dfrac{T_fc^2ZH_{1,\text{inc}}}{2\pi H_0d}\right|^2\right)}\right|^2,$$

where Z is the surface impedance.

Thus,

$$|H_{1,tz}| \sim \frac{\chi T_fc^3|z|^2|H_{1,\text{inc}}|}{2\pi d\left(1+\left|\dfrac{T_fc^2zH_{1,\text{inc}}}{2\pi dH_0}\right|^2\right)} =$$

$$= |H_{1,tz}^{\max}|\cdot\frac{2|H_{1,\text{inc}}/H_{1,\text{inc}}^{\text{opt}}|}{1+|H_{1,\text{inc}}/H_{1,\text{inc}}^{\text{opt}}|^2},\ \text{where}$$

the maximum transmitted magnetic field $H|_{1,tz}^{\max}|$

$$|H_{1,tz}^{\max}| \sim \frac{\chi H_0c|z|}{2}$$

is independent on d, and is obtained when

$$H|_{1,\text{inc}}| = |H_{1,\text{inc}}^{\text{opt}}|H_{1,\text{inc}}^{\text{opt}}| = \frac{2\pi dH_0}{c^2T_f|z|} \sim d.$$

6. THE POLARIZATION OF THE NUCLEI IN FILMS

The formula for the polarization of the nuclei in films (at resonance) is

$$P=\frac{1}{I}\{(I+\tfrac{1}{2})\coth(I+\tfrac{1}{2})\frac{|\alpha|^2}{1+|\alpha|^2}\times$$

$$\times\frac{\cosh[(d-z)/\delta_{\text{eff}}]}{\cosh(d/\delta_{\text{eff}})}\frac{\mu H_0}{kT}-\tfrac{1}{2}\coth\tfrac{1}{2}\frac{|\alpha|^2}{1+|\alpha|^2}\times$$

$$\times\frac{\cosh[(d-z)/\delta_{\text{eff}}]}{\cosh(d/\delta_{\text{eff}})}\frac{\mu H_0}{kT}\},$$

where

$$\alpha=\frac{cT_f[E(0)-E(d)]}{H_0\delta_{\text{eff}}}\coth\frac{d}{\delta_{\text{eff}}},$$

for arbitrary d. If $d<\delta$,

$$\alpha=4\mu T_f H_{1,\text{inc}}/\hbar,$$

i.e. we have OVERHAUSER's formula; if $\delta\ll d\ll\delta_{\text{eff}}$,

$$\alpha=cT_fE(0)/H_0d,$$

and, finally, if $d\gg\delta_{\text{eff}}$, $\alpha=cT_fE(0)/H_0\delta_{\text{eff}}$.

REFERENCES

1. GRISWOLD T. W., KIP A. F., and KITTEL C. *Phys. Rev.* **88,** 951 (1952); FEHER G. and KIP A. F. *ibid.* **95,** 1343 (1954); **98,** 337 (1955).
2. OVERHAUSER A. W. *Phys. Rev.* **89,** 689 (1953); **92,** 411 (1953).
3. CARVER T. R. and SLICHTER C. P. *Phys. Rev.* **92,** 212 (1953).
4. DYSON F. J. *Phys. Rev.* **98,** 349 (1955).
5. SILIN V. P. *Zhurnal eksperimental'noi i teoreticheskoi Fiziki* **30,** 721 (1956).
6. FUCHS, K. *Proc. Camb. Phil. Soc.* **34,** 100 (1938).
7. REUTER G. E. H. and SONDHEIMER E. H., *Proc. Roy. Soc.* **195,** 336 (1948); KAGANOV M. I. and AZBEL' M. YA *Doklady Akademii Nauk SSSR* **102,** 49 (1955); AZBEL' M. YA. and KAGANOV M. I. *ibid.* **95,** 43 (1953); AZBEL' M. YA. *ibid.* **100,** 437 (1955).
8. LIFSHITZ I. M., AZBEL' M. YA., and KAGANOV M. I. *Zhurnal eksperimental'noi i teoreticheskoi Fiziki* **30,** 220 (1956).
9. AZBEL' M. YA. and KANER E. A. *ibid.* **30,** 811 (1956).
10. AZBEL' M. YA., GERASIMENKO V. I. and LIFSHITZ I. M., *ibid,* **31,** 357 (1956).

VOL. 1 1956/5

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 175–187.

SUR LA STRUCTURE DE L'URANIUM MÉTALLIQUE

J. FRIEDEL

Centre de Recherches Métallurgiques de l'École des Mines, de Paris

(*Received* 3 *August*, 1956)

Résumé—L'uranium présente, dans ses trois phases α, β et γ, certaines propriétés physiques caractéristiques d'un métal de transition, avec sans doute un électron par atome, environ, situé dans une bande d'énergie étroite et presque vide, dont nous précisons la forme. Le caractère assez nettement covalent des liaisons interatomiques, surtout en phase α, suggère que 3 à 4 électrons par atome remplissent complètement une seconde bande étroite. Il resterait environ un électron par atome dans une bande de conductibilité de masse normale. Nous proposons d'expliquer les anomalies de hautes températures en phase α par une excitation des électrons de la bande pleine de liaison. Celle-ci serait, en phase α, à environ 0,3 eV au-dessous du niveau de Fermi. Nous soulignons enfin les approximations faites dans cette discussion. Il serait intéressant de confirmer certaines conclusions de ce modèle par des mesures d'effet Hall et d'absorption optique.

Abstract—Various properties of uranium in its α, β and γ phases indicate that it is a transitional metal, with probably about one electron per atom in a narrow and almost empty energy band, the form of which is studied. The strongly covalent character of the interatomic bindings, especially in the α phase, suggests that 3 to 4 electrons per atom fill completely a second narrow band. There should remain about one electron per atom in a broad conductibility band with a normal mass. The full binding band should be at about 0·3 eV below the Fermi level in the α phase; excitation of electrons from this band could explain high temperature anomalies observed in the α phase. The approximations made are discussed. Some conclusions could be checked by measurements of the Hall constant and the optical absorption.

1. L'URANIUM, MÉTAL DE TRANSITION

Il semble établi que l'uranium possède une *forte densité d'états au niveau de Fermi*, c'est-à-dire une forte densité d'états excités d'énergie aussi faible que l'on veut. Ce fait résulte directement des observations suivantes :

Tableau 1. *Susceptibilité molaire et pouvoir thermoélectrique de l'Uranium dans les trois phases.*

phase	α	β	γ
T°K	300	975	1100
χ_A (10^{-6} UEM/mole)	410	480	490
$\frac{d\chi_A}{\chi_A dT}\times 10^4$	0,7	$\simeq 0$	0,6
$\frac{10^3 S}{T}$ (μV/o^2)	26	21	20

(a) La *susceptibilité paramagnétique* de l'uranium est forte dans les trois phases α, β et γ. Elle croît un peu avec la température. Le tableau suivant donne les valeurs de la susceptibilité molaire χ_A, en 10^{-6} UEM/mole, et de $d\chi_A/\chi_A dT$ à trois températures T (d'après Bates et Hughes[1]).

(b) Il possède *à basse température*, donc en phase α, une forte *chaleur spécifique* électronique, proportionnelle à la température (Shoenberg, comm. privée, cf. Goodman et Shoenberg[2]) :

$$C_v = (2\pm 1)\,10^{-3}\,T \text{ cal/o/mole, de 1 à 4°K.}$$

(c) *Au-dessus de* 200°K, la *chaleur spécifique* dépasse la valeur limite $3R \simeq 5{,}95$ cal/o/mole due aux oscillations des ions. La figure 1 donne l'excès $\Delta C_v = C_v - C_v$ ion obtenu à partir des mesures de C_p (d'après Katz et Rabinovitch[3]), en retranchant la petite différence $C_p - C_v$ et la contribution ionique C_v ion. Celle-ci a été calculée par la formule de Debye, avec une

température de DEBYE θ_D raisonnable (170°K). La valeur exacte de θ_D n'a évidemment pas d'importance pour les valeurs de ΔC_v au-dessus de la température ambiante.

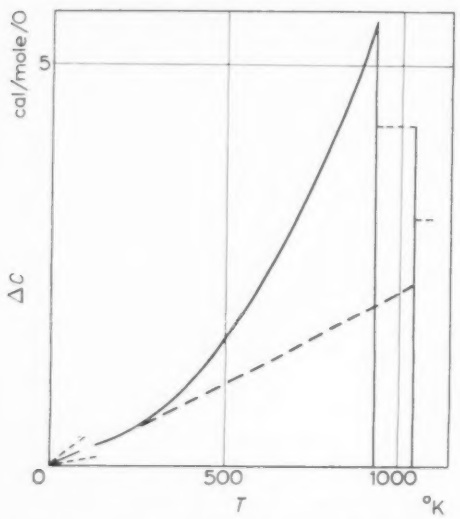


FIG. 1. Anomalie ΔC_v de chaleur spécifique (en cal/o/mole) en fonction de la température. A l'origine, pente initiale, d'après SHOENBERG (les deux flèches marquent les limites extrèmes). En pointillé, extrapolation d'après la formule (2).

Si l'on interprète ces résultats dans l'approximation des bandes, on peut dire que l'uranium possède, dans les trois phases α, β et γ, des structures dont la densité des états est donnée, au voisinage du niveau de Fermi E_M, par les relations (SOMMERFELD et BETHE,[4] STONER[5] :

$$\chi_B = \frac{e^2\hbar^2\Omega}{2m^2c^2} n(E_M) \times$$

$$\times \left[1 + \frac{\pi^2}{6} k^2T^2 \left\{ \frac{d^2n}{n\,d^2E_M} - \left(\frac{dn}{n\,dE_M}\right)^2 \right\} + O_3(T)\right], \quad (1)$$

et

$$\Delta C_v = \frac{2\pi^2\Omega k^2 T}{3} n(E_M) \times$$

$$\times \left[1 + \frac{7\pi^2}{10} k^2T^2 \left\{ \frac{d^2n}{n\,d^2E_M} - \frac{5}{7}\left(\frac{dn}{n\,dE_M}\right)^2 \right\} + O_3(T)\right]. \quad (2)$$

χ_B est ici la susceptibilité molaire dans l'approximation des bandes; n est le nombre d'états par unité de volume et d'énergie; Ω est le volume molaire, et les lettres e, m, $\hbar$, et k ont leur signification usuelle. Enfin on suppose, suivant l'usage, que deux électrons de spins opposés peuvent occuper chaque état.

Ces formules sont valables quand les développements convergent rapidement : à température assez basse et pour des bandes de courbure assez faible pour que les seconds termes des crochets soient petits devant l'unité. Les valeurs de $d\chi_A/\chi_A\,dT$ données ci-dessus montrent que les crochets peuvent être remplacés par l'unité dans la phase α jusqu'à la température ambiante, et aussi dans les phases β et γ jusqu'à plus de 1000°K. On tire ainsi, des valeurs observées pour χ_A et ΔC_v, les valeurs suivantes de $n(E_M)$ exprimées en unités atomiques ($e = \hbar = m = 1$) :

Tableau 2. Nombre $n(E_M)$ d'états par unité de volume et d'énergie, donné par la chaleur spécifique électronique à haut et bas température, et par la susceptibilité paramagnétique. Rapport R entre les valeurs de $n(E_M)$ dèduites du paramagnétisme et de la chaleur spécifique de haute température.

méthode de mesure		α	β	γ
$n(E_M)$	C_v H.T.	0,35	0,75	0,5
	C_v B.T.	0,4±0,2		
	χ	1,15	1,35	1,30
R		3,3	1,8	2,6

Ces valeurs sont fortes, du même ordre que pour le nickel (1,8 U.A. par mesure de C_v), et bien supérieures à celles du cuivre (0,075 U.A.). Comme le nickel, l'uranium est donc dans ses trois phases un *métal de transition* : sa structure de bandes doit présenter, au niveau de FERMI, une bande B, (ou plusieurs) de grande densité $n(E_M)$, donc étroite, ou encore douée d'une forte "masse effective". Une bande parabolique de masse effective μm et contenant p électrons par atome possède en effet au niveau de FERMI une densité d'états

$$n(E_M) = \left(\frac{9\pi p}{4}\right)^{\frac{1}{3}} \frac{\mu}{2\pi^2 r_s} \text{ unités atomiques,} \quad (3)$$

si $v = (4\pi/3)r_s^3$ est le volume atomique. Avec $r_s = 3{,}2$ U.A. et $p \leqslant 6$, le nombre d'électrons extérieurs de l'uranium, il faut une masse effective d'au moins 7 U.A. pour expliquer les densités observés. Ce fait justifie a posteriori que nous ayions négligé dans la formule (1) la contribution *diamagnétique* des électrons de valence. On sait en effet que ce "diamagnétisme de LANDAU" est négligeable dans les bandes à fortes masses effectives. Le diamagnétisme des couches internes est négligeable lui aussi, car certainement inférieur à celui du radon, soit $-24{,}5 \cdot 10^{-6}$ UEM/mole (VAN VLECK[6]).

D'autre part l'accord entre les différentes méthodes de mesure peut être jugé très satisfaisant. Nous rappelons dans un autre article (FRIEDEL[8]) comment le rapport $R = \chi_A/\chi_B$ entre la susceptibilité mesurée χ_A et celle χ_B calculée dans l'approximation des bandes, est lié aux interactions d'échange négligées par cette approximation. Nous y montrons pourquoi R est voisin de 3 dans les métaux de transition (SEEGER[7]).

2. STRUCTURE DE BANDES PRÈS DU NIVEAU DE FERMI

Les propriétés électriques et magnétiques de l'uranium fournissent certains renseignements sur ce point.

2.1. *Variation du Paramagnétisme avec la Température*

Un certain nombre de métaux de transition possèdent, comme l'uranium, un paramagnétisme qui croît avec la température. Un tel comportement s'explique, suivant la formule (1), par une *courbure positive* d^2n/dE_M^2 de la structure de bandes au niveau de Fermi (STONER,[9] WOHLFARTH,[10] KRIESSMAN et CALLEN[11]). Dans l'application de la formule (1), il faut tenir compte de la différence entre les susceptibilités mesurée χ_A et calculée χ_B. On a

$$\frac{d\chi_A}{\chi_A\, dT} = R\frac{d\chi_B}{\chi_B\, dT}.$$

Le tableau suivant donne les valeurs obtenues pour $(d^2n/ndE_M^2) - (dn/n\,dE_M)^2$, exprimées en unités atomiques, à partir des valeurs de $d\chi_A/\chi_A dT$ et de R données précédemment.

Tableau 3. Les valeurs de la courboure de bande au niveau de Fermi, données par la variation du paramagnétisme avec la température.

Phases	α	β	γ
$\frac{d^2n}{n\,dE_M^2} - \left(\frac{dn}{n\,dE_M}\right)^2$	$0{,}25 \cdot 10^4$	$\simeq 0$	$0{,}05 \cdot 10^4$
$\frac{dn}{n\,dE_M}$ d'après (5)	-5	-4	$-3{,}5$
$\frac{dn_1}{n_1\,dE_M}$ d'après (6)	29	23,5	22

2.2. *Pouvoir Thermoélectrique*

L'uranium est doué d'un *fort* pouvoir thermoélectrique S *positif*. Dans l'approximation des bandes, ce fait indique que la résistivité $\rho(E)$ offerte par l'agitation thermique aux électrons d'énergie E varie très rapidement avec E au voisinage du niveau de FERMI. On a en effet, aux températures moyennes (supérieures à la températures de DEBYE, mais pas trop élevées, cf. MOTT et JONES[12]),

$$S = \frac{\pi^2}{3}\frac{k^2T}{e}\frac{d\ln\rho}{dE_M}. \quad (4)$$

Cette variation rapide, comme d'ailleurs les fortes valeurs observées pour la résistivité $\rho(E_M)$ elle-même, sont certainement liées à la présence d'une bande étroite au niveau de FERMI. Mais les valeurs de S peuvent être interprétées de deux façons différentes suivant que le métal possède ou non une "bande de conductibilité" de masse effective voisine de la masse électronique normale.

(1) Pas de bande de conductibilité. Si la bande étroite B_1 chevauche seule le niveau de FERMI, on peut écrire

$$S = -C\frac{\pi^2}{3}\frac{k^2T}{e}\frac{dn}{n\,dE_M}, \quad (5)$$

où le coefficient C est voisin de 6 pour une bande parabolique (cf. MOTT et JONES[12]). Une valeur

positive de S correspondrait ainsi à une bande B_1 à pente dn/dE_M *négative* au niveau de FERMI.

(2) Une bande de conductibilité. Comme l'a remarqué MOTT pour les éléments de transition usuels (cf. MOTT et JONES[12]), la situation est différente si, en plus de la bande étroite B_1, le niveau de FERMI est chevauché par une bande B_2 plus large, donc de "masse effective" plus voisine de la masse normale, ou encore possédant un "nombre effectif" de porteurs plus grand que B_1. B_2 sera en effet la bande de conductibilité : ses électrons, plus mobiles, transporteront la majeure partie du courant électrique; la résistivité sera, en majeure partie, proportionnelle à leur probabilité de transition au niveau de FERMI vers la bande étroite B_1, donc à sa densité $n_1(E_M) \simeq n(E_M)$ au niveau de FERMI. L'équation (4) s'écrit dans ce cas

$$S \simeq \frac{\pi^2}{3}\frac{k^2T}{e}\frac{dn_1}{n_1 dE_M}. \qquad (6)$$

Un fort pouvoir thermoélectrique positif correspond dans ce cas à une bande étroite B_1 à pente dn_1/dE_M *positive*.

Les équations (5) et (6) sont évidemment très approchées; quand plusieurs bandes étroites B_1 chevauchent le niveau de FERMI, elles doivent s'appliquer à la densité totale de ces bandes. Le premier tableau donne les valeurs de S (en μV/o) observées dans les trois phases (d'après THOMPSON,[13] et DAHL et VAN DUSEN[14]). Le troisième tableau donne les valeurs, en unités atomiques, qu'on en déduit pour $dn/n\,dE_M$ d'après (5) et $dn_1/n_1\,dE_M$ d'après (6).

Les figures 2 et 3 rassemblent graphiquement les résultats ainsi obtenus sur la structure $n_1(E)$ de

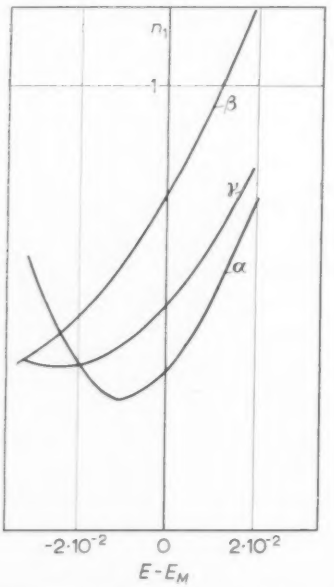


FIG. 2. Structure $n_1(E)$ la bande étroite B_1 dans les trois phases α, β et γ (en unités atomiques), s'il n'y a pas de bande de conductibilité de masse normale.

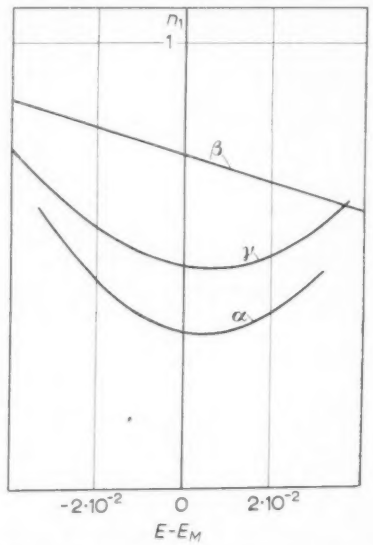


FIG. 3. Comme fig. 2, s'il y a une bande de conductibilité.

la bande étroite B_1 au voisinage du niveau de FERMI E_M, dans les deux cas envisagés et pour les trois phases α, β et γ. Des mesures d'effet Hall permettraient de trancher entre les deux cas possibles pour chacune des phases. A notre connaissance, de telles mesures n'ont pas encore été faites. La nature des liaisons dans l'uranium, que nous allons étudier, ainsi que l'analyse faite au §4 des anomalies observées en phase α, semblent en meilleur accord avec les structures de la figure 3, qui supposent *l'existence d'une bande de conductibilité B_2 de masse normale.*

3. NATURE DE LA BANDE ÉTROITE

L'atome d'uranium à l'état gazeux a probablement la configuration $5f^36d7s^2$ (KATZ et

RABINOVITCH[3]). Il est donc vraisemblable que, dans l'état métallique, ses électrons de valence conservent un caractère 5*f* et 6*d* marqué. Nous étudions à ce point de vue la structure cristalline de l'uranium et l'aspect des spectres de rayons X.

3.1. *Nature des Liaisons dans l'Uranium α*

La forte courbure positive $d^2n/d\,E_M{}^2$ de la structure de bandes an niveau de FERMI E_M s'explique *si deux bandes étroites se chevauchent à son voisinage.* La figure 4(a) schématise la structure probable de ces bandes, telle qu'on peut la déduire de la nature des liaisons : la bande étroite presque pleine B_1 contient 4 électrons par atome; elle est construite sur des *orbitales de liaison hybrides, à caractère surtout* 5*f* et 6*d. Les orbitales* 7*s donnent lieu à une bande de conductibilité* B_2, de masse voisine de la masse normale. Cette bande se partage avec la bande étroite presque vide B'_1 les deux électrons de valence par atome restants. Nous examinons ces différents points successivement.

3.1.1. *Liaisons covalentes.* La structure orthorhombique de l'uranium, rappelée figure 5, est

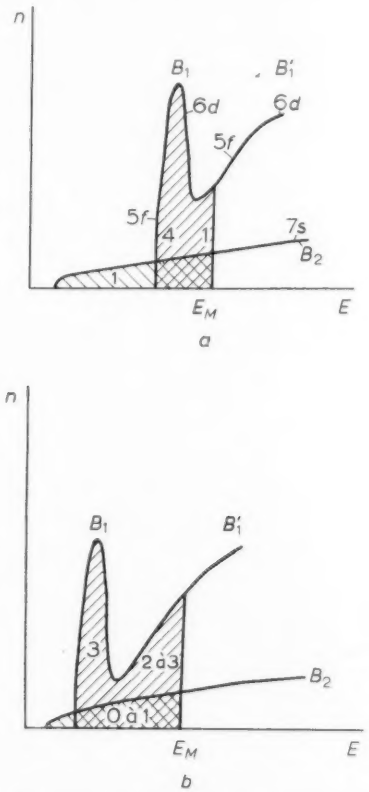


FIG. 4. Structure de bandes proposée pour l'uranium. a—Phase α; b—Phase γ. Les chiffres cerclés donnent le nombre d'éleetrons par atome dans les trois bandes B_1, B'_1 et B_2.

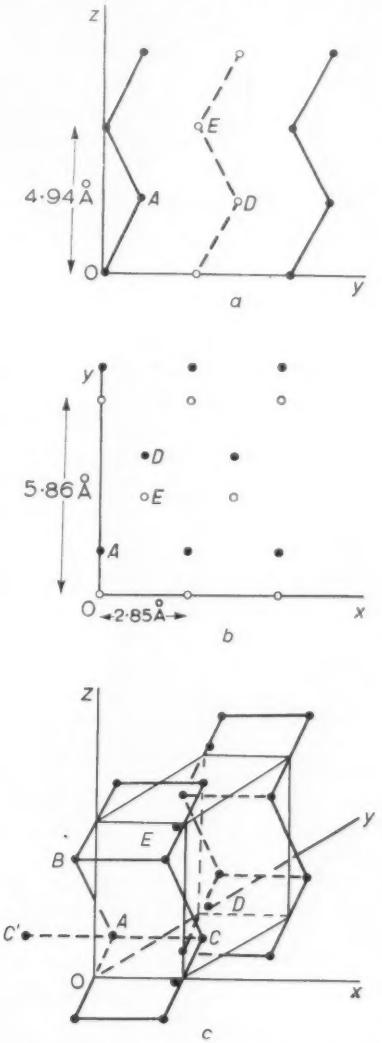


FIG. 5. Structure de l'uranium α : a — coupe yOz; b — coupe xOy; c — perspective cavalière. Les points représentent des atomes dans le plan considéré; les cercles, des atomes dans le plan cristallographique parallèle voison.

tout à fait particulière. Comportant deux atomes par maille, elle est constituée par des files d'atomes, parallèles à un axe OZ, un peu en zig zag dans des plans parallèles à yOz et plantées en quinconce sur un plan xOy. La distance entre files, de 2,85 Å suivant Ox, est peu supérieure à la distance interatomique dans les files, soit 2,75 Å. Les files forment ainsi de véritables tôles ondulées parallèles à xOz. Ces tôles, empilées les unes sur les autres, sont maintenues ensemble de façon plus lâche, car les distances entre atomes de tôles voisines ne sont pas inférieures à 3,25 Å. Les 12 atomes premiers voisins d'un atome donné se répartissent ainsi de la façon suivante :

2 atomes à 2,75 Å (liaisons AB, figure 4)
2 atomes à 2,85 Å „ AC „
4 atomes à 3,25 Å „ AD „
4 atomes à 3,34 Å „ AE „

Cette structure est très différente d'un empilement compact. Elle présente un caractère *covalent marqué et très particulier* (PAULING[15]) : chaque atome forme 4 liaisons (AB, AB', AC, AC') *à peu près coplanaires et à* 90° *les unes des autres*. Les zig-zags des files font que les deux liaisons opposées de type AB, AB' forment en fait entre elles un angle inférieur à 180° (environ 125°).

3.1.2. *Bandes de liaison et d'antiliaison.* Quand on réunit des molécules pour former un solide, l'énergie nécessaire pour rompre une de leurs liaisons covalentes* n'est plus parfaitement déterminée : elle dépend de la vitesse à laquelle la liaison rompue passe d'une molécule à sa voisine. On peut dire que le niveau d'énergie de la liaison s'élargit en une "bande de liaison". La rupture d'une liaison correspond à l'introduction d'un trou positif dans la bande; son énergie dépend un peu de sa vitesse, ou de son "moment".

On sait d'autre part que, dans l'approximation des orbitales moléculaires, les électrons de liaison sont supposés être dans des orbitales de liaison $\psi_L \simeq \psi_A + \psi_{A'}$ égales en première approximation à une combinaison linéaire d'orbitales atomiques ψ_A et $\psi_{A'}$. La combinaison inverse $\psi_{L'} = \psi_A - \psi_A$ correspond approximativement à un état excité, dit d'antiliaison. Quand on passe d'une molécule à un solide, l'énergie nécessaire pour amener (de l'infini) un électron dans un tel état s'élargit également en une *bande d'antiliaison* d'énergie moyenne supérieure à celle de la bande de liaison.

Les liaisons covalentes nettes de l'uranium α indiquent que les bandes de liaison et d'antiliaison correspondantes sont bien marquées, distinctes l'une de l'autre, et que la première est à peu près pleine tandis que la seconde est à peu près vide. La bande de liaison est évidemment la bande B_1 de la figure 4, la bande d'antiliaison fait partie de la bande B'_1.

3.1.3. *Caractère des liaisons.* Nous voulons montrer maintenant que les liaisons s'obtiennent par hybridisation d'orbitales $5f$ et $6d$.

L'existence, dans certains solides, de liaisons *fortes et asymétriques* est en général mise en rapport avec un caractère p marqué des fonctions d'onde des électrons de valence. Les orbitales p sont en effet *antisymétriques*, c'est-à-dire qu'elles prennent des valeurs égales et des signes opposés en deux points symétriques par rapport au noyeau. Par combinaison linéaire avec des orbitales symétriques s ou d, elles donnent ainsi des orbitales atomiques hybrides ψ_A fortement dirigées dans une certaine direction de l'espace. Les orbitales hybrides construites uniquement avec des orbitales s et d sont au contraire dirigées également dans deux directions opposées. Les obitales du premier type, schématisées figure 6(a), ont évidemment, toutes choses égales d'ailleurs, un centre de gravité plus éloigné

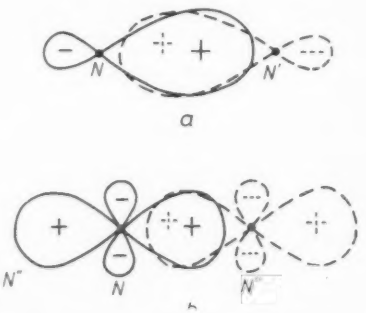


FIG. 6. Caractère des orbitales hybrides de liaison ψ_A : a — contenant, entre autres, des orbitales p ou f; b — contenant uniquement des orbitales s et d. N dénote le noyeau; les signes + et −, ceux de la fonction d'onde. En pointillé, l'orbitale de liaison $\psi_{A'}$.

* En envoyant l'électron correspondant à l'infini hors du solide.

VOL. 1 1956/57

du noyau N que celles du second type, figure 6(b). Elles peuvent recouvrir plus largement une orbitale $\psi_{A'}$ du même type, centrée sur un noyau N' voisin et dirigée en sens inverse; elles forment donc avec elles des orbitales moléculaires $\psi_L = \psi_A + \psi_{A'}$ conduisant à des liaisons covalentes plus fortes. Pour profiter pleinement de ses possibilités, une orbitale symétrique du type b doit de plus former des liaisons avec au moins les deux atomes voisins N'' situés dans des directions opposées; les liaisons asymétriques du type a peuvent au contraire s'ordonner autour d'un noyeau N sans l'admettre comme centre de symétrie. Le tétraèdre du carbone tétravalent fournit un exemple classique de liaison du type a ($2p$); la liaison hydrogène ($1s$), les "complexes" chimiques formés avec des métaux de transition comme le cobalt ($3d$) et certaines structures métalliques comme celle du chrome cubique centré ($3d$) sont des exemples au moins approximatifs du type b.*

Dans l'uranium α, les orbitales de valence antisymétriques les plus stables ont un caractère $5f$ plutôt que $7p$. Nous pouvons donc admettre que les quatre liaisons covalentes AB, AB', AC, AC' ont un fort caractère $5f$. Un examen des orbitales possibles $7s$, $6d$ et $5f$ montre que les orbitales hybrides nécessaires pour former quatre liaisons coplanaires, parallèles à Ox et à Oz, sont approximativement des combinaisons linéaires de quatre types :

$$\psi_{x\pm} = a_s + a_d(2x^2-y^2-z^2) \pm \pm[a_{f_1}(x^2-3z^2)x + a_{f_2}(x^2+z^2-4y^2)x],$$

$$\psi_{z\pm} = a'_s + a'_d(2z^2-x^2-y^2) \pm \pm[a_{f_1}(z^2-3x^2)z + a'_{f_2}(x^2+z^2-4y^2)z].$$

Les deux fonctions a_d et a_{f_1} de la distance $r = (x^2+y^2+z^2)^{\frac{1}{2}}$ au noyeau doivent avoir des amplitudes nettement supérieures aux deux autres a_s et a_{f_2}, car les termes correspondant donnent des orbitales hybrides plus nettement dirigées. Les termes nécessaires pour expliquer que les *liaisons* ne sont pas tout à fait coplanaires sont sans doute petits également.

* L'importance dans les solides de cette distinction entre liaisons covalentes symétriques et asymétriques a été soulignée récemment par DEHLINGER.[16]

3.1.4. *Structure des bandes.* La discussion qui précède indique que, quand on forme la phase α, la bande B_1 a vraisemblablement un caractère mixte $5f$-$6d$. Cette *hybridisation* est caractéristique des bandes de liaison et d'antiliaison (cf. par exemple KIMBALL[17] pour les bandes $2sp$ du diamant). La figure 7 montre schématiquement

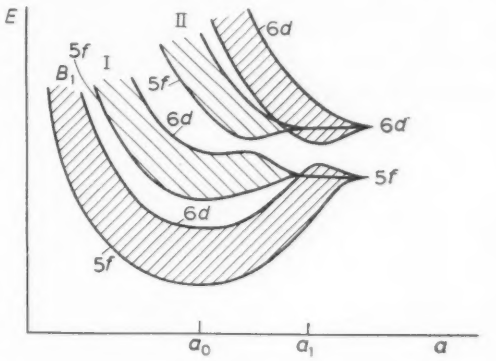


FIG. 7. Nature présumée des bandes étroites B_1, B'_1 dans l'uranium α, en fonction du paramètre cristallin pour une direction donnée dans l'espace des moments. Les zones d'énergie permises sont hâchurées; a_0 est la valeur du paramètre observé.

qu'une forte hybridisation des bandes se produit nécessairement pour des paramètres cristallins inférieurs à une valeur critique a_1 du fait que les bandes issues des niveaux atomiques voisins $5f$ et $6d$ ne peuvent, en général, se couper.

Les états $5f$ et $6d$ qui ne sont pas utilisés dans la bande de liaison B_1 et dans la bande d'antiliaison correspondante donnent évidemment des bandes qui ne s'élargissent que pour des paramètres cristallins plus petits. Ces bandes s'hybridisent moins fortement et ont des énergies qui doivent être comprises entre celles des bandes de liaison et d'antiliaison. Elles sont schématisées figure 7 par les bandes intermédiaires I, II. La bande B'_1 sera évidemment celle de ces bandes intermédiaires dont l'énergie est la plus basse (bande I, figure 7). La variation des limites de bandes avec la direction de moment considérée permet le petit chevauchement observé des bandes B_1 et B'_1.

On peut tirer les conséquences suivantes de l'hybridisation $5f\,6d$:

(1) La *bande de liaison* B_1 *doit être étroite* : sa largeur doit être de l'ordre de l'énergie d'excita-

tion $6d$–$5f$ dans l'atome libre,* et cette énergie est mal connue, mais certainement faible, inférieure à l'énergie d'ionisation de l'uranium, soit 4 eV (KIESS, HUMPHREYS et LAUN[18]).

(2) La *bande* B_1 *contient autant d'électrons que de liaisons*, soit *quatre par par atome* (cf. Annexe).

(3) Comme B_1 contient vraisemblablement peu d'états $7s$, il doit s'y superposer *une bande* B_2 *contenant surtout des états* $7s$. Comme dans tous les solides, cette bande doit être large, ou encore avoir une masse effective μm voisine de la masse normale† m.

(4) C'est donc la figure 3 qui s'applique, plutôt que la figure 2. L'aire comprise entre le niveau de FERMI et le minimum de la bande étroite indique qu'il doit y avoir environ un électron par atome dans la bande B'_1, donc $6-4-1=1$ électron par atome environ ‡ dans la bande B_2 (figure 4a).

* La largeur de la bande provient en effet d'un résonance entre liaisons. Le mélange d'orbitales atomiques qui en résulte détruit, dans une certaine mesure, leur hybridisation $5f$–$6d$, en redonnant des fonction plus voisines des orbitales composantes $5f$ et $6d$. Contrairement à l'uranium, le carbone (diamant ou graphite) a une bande de liaison large, car l'énergie d'excitation $2s$–$2p$ correspondante est assez forte : environ 3 eV de $2s^2 2p^2$ à $2s2p^3$ pour l'atome libre (MOORE[19]).

† On sait en effet qu'au voisinage du centre de la zone, et en supposant la bande isotope, on a, en unités atomiques,

$$\frac{1}{\mu}=1+2\sum_{i\neq 7s}\frac{|\int\psi_{7s}{}^{*}\operatorname{grad}\psi_i\,d\tau|^2}{E_{7s}-E_i},$$

où ψ_i est la fonction de BLOCH d'énergie E_i et de moment nul. Les ψ_i sont evidemment des combinaisons linéaires de fonctions atomiques autres que $7s$. Seules les fonctions p donnent des termes non nuls dans la somme; et comme $7p$ a une énergie bien supérieure à $7s$, μ doit être voisin de l'unité.

‡ Le caractère faiblement $7s$ de la bande B_1 correspond évidemment à une hybridisation qui se produit dans la zône d'énergies où la bande B_2 rencontre la bande B_1. L'hybridisation vient de ce que les deux bandes ne peuvent pas se couper. Comme la bande B_2 est large, le petit mélange se produit dans une région restreinte des énergies et des moments, qui n'intéresse pas ou peu le niveau de FERMI. Un mélange analogue se produit pour les bandes d des éléments de transition, et n'empêche également pas les bandes d et s d'avoir un sens physique au niveau de FERMI (cf. KRUTTER,[20] SLATER et KOSTER[21]).

3.2 *Spectres de Rayons X*

Par analogie avec ce qui se passe pour les atomes isolés nous avons supposé, dans la figure 7, le niveau atomique $5f$ plus stable que le niveau $6d$. Nous en avons déduit que les bandes B_1 et B'_1 avaient un caractère surtout $5f$ dans les basses énergies, surtout $6d$ dans les hautes énergies. Cette conclusion est en accord avec la forme des spectres d'absorption $L_{II,\,III}$ et $M_{IV,\,V}$, c'est-à-dire ceux où un trou positif est créé dans un état $2p$ et $3d$ respectivement. Dans une absorption $L_{II,\,III}$, l'intensité est essentiellement proportionnelle à la densité des états finaux à caractère d. Or on observe une "raie blanche", c'est à dire un fort maximum d'intensité, à environ une dizaine d'eV au-dessus de la limite d'absorption (KATZ et RABINOVITCH[3]). Cette raie blanche correspondrait, d'après les figures 4 et 7, à des transitions vers le haut d'une bande d'énergie supérieure (telle que la bande I, figure 7).

L'absorption $M_{IV,\,V}$, qui mesure essentiellement la densité des états à caractère f, est au contraire importante dès la limite d'absorption; ce fait est en accord avec le caractère $5f$ du bas de la bande B'_1, figures 4 et 7, où est supposé se situer le niveau de FERMI E_M. En d'autres termes, en calculant la différence, représentée figure 8,

$$\delta=(M_{IV,\,V})_{abs}-\{L_{III\,abs}+(M_{IV,\,V}L_{III})_{em}\},\quad(7)$$

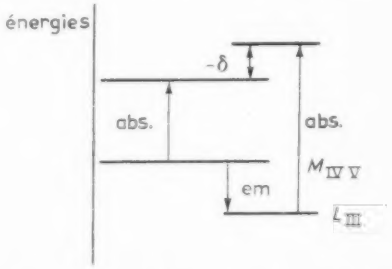


FIG. 8. Comparaison des absorptions L_{III} et $M_{VI,\,V}$ dans l'uranium.

et en comptant l'absorption L_{III} *jusqu'à la raie blanche*, on trouve une quantité *négative*, -7 eV d'après CAUCHOIS.[22] δ correspond évidemment à la différence d'énergie entre la limite d'absorption et la raie blanche de L_{III}.

L'anomalie δ calculée par l'équation (7) est représentée, figure 9, en fonction du nombre

atomique de l'élément considéré (d'après CAUCHOIS[22]). On voit qu'avant le radium elle est positive, en accord avec le fait que pour ces éléments les états atomiques 5*f* sont instables, et que

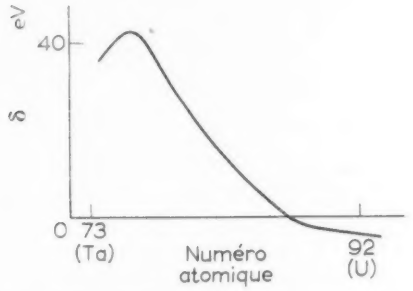


FIG. 9. Variation de l'anomalie δ d'absorption $M_{IV,V}$ en fonction du nombre atomique.

la bande est à caractère *d* au niveau de Fermi E_M, donnant une forte absorption L_{III} dès l'énergie E_M; δ doit donc mesurer directement, dans ce cas, la différence d'énergie entre E_M et une bande 5*f*. C'est ce que confirme un calcul grossier (FRIEDEL[8]). Le changement de signe de δ pour le radium, prouve donc assez directement que les éléments plus lourds ont, dans l'état métallique, *des électrons de valence à caractère 5f marqué.*

3.3 *Nature des Liaisons dans l'Uranium β et γ*

La phase γ a la structure cubique centrée. Elle se range donc dans la même série que les éléments de transition de même structure, qui ont en général 5 ou 6 électrons extérieurs (Ti, V, Cr, Mn, Fe, Zr, Nb, Mo, Ta, W). Ce fait suggère, comme dans ces éléments, des liaisons à caractère essentiellement *d*.

La discussion du §3.1.3. montre en effet que le grand nombre de liaisons que chaque atome doit former (8 avec des atomes premiers voisins, 6 avec des atomes seconds voisins à distance peu supérieure), et leur disposition symétrique par rapport à son noyeau, favorisent les liaisons covalentes symétriques décrites figure 6(b), plutôt que les liaisons asymétriques de la figure 6(a), donc ici les orbitales 6*d* et 7*s* plutôt que 5*f*.

Nous pouvons donc admettre, comme dans le chrome, l'existence d'une bande *d* étroite et pleine B_1, construite sur les orbitales *d* orthogonales de types

$$yz+zx+xy$$

$$yz-zx+xy$$

$$yz-zx-xy.$$

L'intensité de ces orbitales est en effet dirigée parallèlement aux quatre diagonales du cube $x = \pm y = \pm z$. Elles permettent donc de former les 8 liaisons entre atomes premiers voisins (cf. GANZHORN,[23] SEEGER,[7] SLATER et KOSTER[21]). Nous rappelons dans l'annexe comment cette bande doit contenir autant d'électrons que d'orbitales *d* indépendantes, soit *trois par atome.* Les deux orbitales *d* restantes, de types z^2-y^2 et z^2-x^2, sont moins stables, car elles ne permettent de faire que des liaisons entre seconds voisins. Elles se mélangent sans doute avec des états 5*f* pour former une seconde bande étroite B'_1 incomplètement occupée.*

Enfin, comme en phase α et pour la même raison, les orbitales 7*s* doivent donner lieu à une bande de conductibilité B_2 large, de masse normale.

Nous sommes ainsi conduit à proposer la structure de bandes de la figure 4(b). Cette structure est en accord avec celle de la figure 3, si nous supposons que la bande B'_1 contient entre 3 et 2 électrons par atome, donc la bande B_2 entre 0 et 1 électron par atome. Ces chiffres semblent raisonnables. Il serait intéressant de confirmer le nombre d'électrons situés dans la bande de conductibilité B_2 par des mesures d'effet Hall. S'il n'y avait pas de bande de conductibilité, il serait au contraire très difficile d'imaginer une bande ayant l'allure donnée figure 2 et néanmoins assez étroite pour ne contenir que 6 électrons par atome.

La structure β présente également des liaisons covalentes nettes; les électrons correspondants sont dans une bande étroite pleine, analogue à la bande B_1, figure 4. Le nombre de liaisons par atome est réduit (2, 3 ou 4).

Certaines de ces liaisons sont du type "symétrique", figure 6(b); les électrons de liaison sont donc en nombre inférieur à celui des liaisons (cf. Annexe) : il y a moins de 4 électrons de liaison par atome dans la bande B_1. La figure 3 en donne 2 environ dans la bande presque vide B'_1. Il en

* La bande d'antiliaison correspondant à B_1 a sans doute une énergie nettement supérieure (cf. SEEGER[7]).

reste donc environ 1 pour la bande de conductibilité B_2. L'importance des orbitales $5f$ dans les liaisons est soulignée par la complexité de la structure.

4. ANOMALIES DE HAUTE TEMPÉRATURE DANS LA PHASE α

4.1 *Nature des anomalies observées*

La figure 1 montre que l'excès ΔC_v de chaleur spécifique croît rapidement avec la température en haut de la phase α, et de façon à peu près exponentielle. La courbe en pointillé montre que le terme en T^3 dans la formule (2) est insuffisant pour expliquer la courbure observée. La dilatation linéaire mesurée aux rayons X, et la résistivité électrique présentent des anomalies exponentielles similaires (figure 10 et 11, d'après *The*

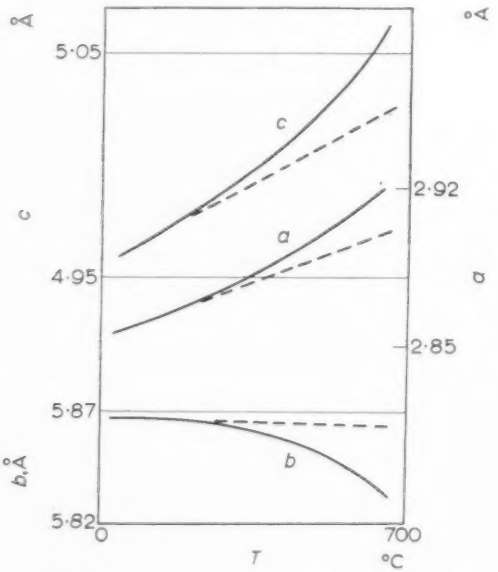


FIG. 10. Variation avec la température T des paramètres cristallins a, b, c (suivant les axes $Oxyz$ de la figure 5), d'après *The Reactor Handbook*.[24] En pointillé, extrapolation de la variation linéaire de basse température.

Reactor Handbook,[24] DAHL et VAN DUSEN[14]). L'anomalie négative de la résistivité ρ est spécialement surprenante. Si l'on suppose en effet ρ due à l'agitation thermique et proportionnelle à la densité des états dans la bande étroite au niveau de Fermi, on devrait avoir (MOTT et JONES[12]) :

$$\rho = \text{const}\, T\left[1+2\alpha\gamma T+ \right.$$

$$\left. +\frac{\pi^2}{6}k^2T^2\left\{\frac{d^2n_1}{n_1\,dE_M{}^2}-2\left(\frac{dn_1}{n_1\,dE_M}\right)^2\right\}\right]+O_4(T). \tag{8}$$

Dans cette formule, le terme en T^2 est dû à la variation avec T de la température de DEBYE; le

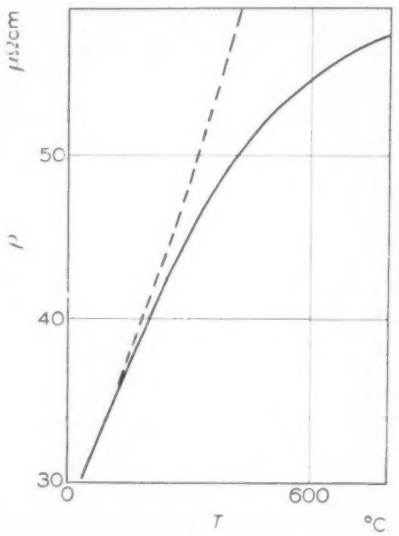


FIG. 11. Résistivité de l'uranium α en fonction de la température T (d'après DAHL et VAN DUSEN[14]). En pointillé extrapolation à l'aide de la formule 8.

terme en T^3 à l'élargissement avec T du niveau de Fermi. α est le coefficient de dilatation thermique; γ le coefficient de GRÜNEISEN. On a $\alpha\gamma \simeq 10^{-4}$ C.G.S. (KATZ et RABINOVITCH[3]); et, avec la valeur positive de la courbure $d^2n_1/dE_M{}^2$ calculée précédemment, la formule (8) donne pour $\rho(T)$ une très légère courbure positive, suivant la courbe pointillée de la figure 11.

4.2. *Analyse des anomalies*

Des anomalies analogues sont observées dans les métaux alcalins près du point de fusion (cf. MACDONALD[25] par exemple). Elles sont expliquées, dans ce cas, par la présence en équilibre à haute température de *défauts* ponctuels (lacunes), d'énergie U assez faible pour être en quantité notable.

VOL. 1 1956/5

En écrivant ici que les anomalies des paramètres cristallins a, b, c et de la résistivité ρ sont proportionnelles à la concentration

$$C = ne^{-U/kT} \tag{9}$$

de défauts dont la nature reste à préciser, on trouve les valeurs suivantes pour U (en eV) et pour les anomalies relatives par nombre n de défauts.

Tableau 4. *Les valeurs pour* U (en eV) *et pour les anomalies de dilatation thermique et de la résistivité électrique.*

Grandeur physique	Anomalie relative pour n défauts	U(eV)
a	$\Delta a/a = 0{,}67$	0,35
b	$\Delta b/b = -0{,}46$	0,36
c	$\Delta c/c = 0{,}71$	0,36
ρ	$\Delta\rho/\rho \simeq -3$	$\simeq 0{,}2$
C_v	$n = 4{,}0$	0,28

Nous avons également analysé l'anomalie de chaleur spécifique (différence entre les courbes continue et pointillée de la figure 1), d'après l'équation

$$\Delta C_v = \frac{\partial}{\partial T}(NUC) = \frac{nNU^2}{kT^2} e^{-U/kT}, \tag{10}$$

où N est le nombre d'AVOGADRO. Nous avons ainsi obtenu les valeurs $U = 0{,}28$ eV et $n = 4{,}0$ portées au tableau.

4.3 *Nature des défauts*

Les énergies de création U ainsi obtenues sont du même ordre de grandeur et faibles. Il est très peu vraisemblable que ces défauts soient des lacunes : par suite du caractère fortement covalent des liaisons, l'énergie des lacunes doit être peu inférieure a l'énergie de sublimation, soit plus de 4 eV (KATZ et RABINOVITCH[3]).

Deux autres explications sont possibles, d'ordre l'une atomique et l'autre électronique; nous allons montrer que le seconde semble plus satisfaisante.

4.3.1. *Atomes interstitiels.* Dans une structure assez peu dense comme l'uranium α, l'introduction de tels atomes demande peut-être une énergie élastique assez réduite. Comme les liaisons ne sont pas saturées, ces atomes pourraient peut-être récupérer une partie de leur énergie de distorsion en formant des liaisons covalentes avec les atomes voisins. Les énergies U observées semblent cependant faibles, même pour un tel modèle : le métal a de fortes constantes élastiques; il n'y a pas de position favorable évidente pour les atomes interstitiels; enfin on voit mal comment leur introduction peut produire une contraction suivant l'axe Oy normal au plan des "tôles ondulées". D'ailleurs, s'ils avaient des positions stables, les atomes interstitiels devraient diffuser lentement. Les anomalies observées devraient présenter une forte hystérésis à température variable. Il serait intéressant d'examiner si une telle hystérésis existe, car elle ne pourrait être attribuée aux défauts que nous allons maintenant discuter.

4.3.2. *Rupture de liaisons covalentes.* Il est plus naturel d'expliquer les anomalies observées par une rupture des liaisons covalentes, c'est à dire une excitation des électrons de la bande de liaison B_1 dans la bande B'_1, figure 4(a), sous l'action de la température. La forte courbure d^2n_1/dE_M^2 de la bande étroit au niveau de FERMI fait en effet que les développements (1), (2) et (8) ne sont certainement pas valables à haute température. Les énergies U obtenues indiquent que la position moyenne de la bande B_1 est à environ 0,3 eV $\simeq 10^{-2}$ U.A. au-dessous du niveau de FERMI, une valeur très raisonnable d'après la figure 3. La bande B_1 a d'ailleurs sans doute une certaine largeur. Ainsi peut s'expliquer que des phénomènes différents, qui font des moyennes différentes sur les états de cette bande, conduisent à des valeurs de U un peu différentes. Deux autres points sont encore en accord avec cette explication : le nombre $n = 4$ de défauts par atome, et l'expansion qu'ils produisent dans les plans des tôles ondulées, c'est-à-dire parallèlement aux liaisons.

La contraction suivant b peut alors s'expliquer en termes élastiques comme une contraction de POISSON. En termes électroniques, on peut distinguer une plus forte attraction par les noyeaux des électrons autres que ceux de liaison (il y a moins d'électrons de liaison pour leur faire écran), et peut-être la formation de liaisons entre atomes

de "tôles ondulées" voisines. L'anomalie négative de résistivité observée traduit dans doute cette consolidation entre tôles, qui doit réduirel' agitation thermique en empêchant les tôles de vibrer indépendamment les unes des autres. Cette élévation de la "température de DEBYE" moyenne est sans doute le facteur de beaucoup le plus important : l'augmentation avec T du nombre effectif de porteurs de charge doit être minime, car les trous introduits dans la bande de liaison sont sûrement peu mobiles.*

En conclusion les anomalies observées s'expliquent assez bien par l'excitation, à haute température, d'électrons de la bande de liaison B_1. La faiblesse des énergies de formation mises en jeu correspond à des concentrations importantes de trous positifs créés dans B_1, plusieurs % par atome à la température de transformation.

L'auteur tient à remercier les Drs. J. BLIN, H. CURIEN, M. HERPIN, J. PÉRIO et W. M. LOMER pour leurs critiques pertinentes.

ANNEXE—

Bande de liaison dans l'approximation des liaisons fortes

Dire qu'un cristal possède des liaisons covalentes assez marquées, c'est dire que certains de ses électrons ont des fonctions d'ondes voisines de celles ψ_i de ces liaisons. Dans l'approximation des bandes, où chaque électron est supposé distribué également sur toutes les mailles du réseau cristallin, les fonctions d'ondes ψ de ces électrons doivent donc être des combinaisons linéaires de ψ_i. La combinaison linéaire de moment $\boldsymbol{k}$ qui satisfait aux conditions cycliques de BORN s'écrit ainsi, au premier ordre (approximation dite des liaisons fortes):

$$\psi(\boldsymbol{r}) = \sum_j e^{i\boldsymbol{k}\boldsymbol{R}_j} \sum_i a_i \psi_i(\boldsymbol{r}-\boldsymbol{R}_j). \tag{11}$$

Dans ce développement, la sommation j se fait sur les vecteurs $\boldsymbol{R}_j$ du réseau de Bravais du cristal; celle i sur les diverses liaisons covalentes d'une maille du cristal. Les ψ_i sont les orbitales (moléculaires) correspondant à ces liaisons; elles sont solutions d'équations du type

$$H_i\psi_i \equiv [-\tfrac{1}{2}\Delta + V_i(\boldsymbol{r}-\boldsymbol{R}_d)]\psi_i(\boldsymbol{r}-\boldsymbol{R}_j) = E_i\psi_i(\boldsymbol{r}-\boldsymbol{R}_j). \tag{12}$$

V_i est ici le potentiel du couple d'atomes formant la liaison i; E_i est l'énergie de liaison quand le couple d'atomes est isolé dans le vide (si l'on se fixe la condition $\psi_i(\infty) = 0$).

Les fonctions ψ satisfont à l'équation

$$H\psi \equiv [-\tfrac{1}{2}\Delta + V_R(\boldsymbol{r})]\psi = E\psi, \tag{13}$$

où V_R est le potentiel périodique du réseau. En multipliant cette équation à gauche par $\psi_{i'}{}^*(\boldsymbol{r})$, en intégrant et en éliminant les coefficients a_i du système d'équations linéaires ainsi obtenues, on obtient, suivant le processus classique, l'équation suivante en E en fonction du moment $\boldsymbol{k}$:

$$|(E-E_i)X_{ii'} - Y_{ii'}| = 0$$

$$\text{avec } X_{ii'} = \sum_j e^{i\boldsymbol{k}\boldsymbol{R}_j} \int \psi_{i'}(\boldsymbol{r})\psi_i(\boldsymbol{r}-\boldsymbol{R}_j)\, d\tau \tag{14}$$

$$Y_{ii'} = \sum_j e^{i\boldsymbol{k}\boldsymbol{R}_j} \times \int \psi_{i'}{}^*(\boldsymbol{r})[V_R(\boldsymbol{r}) - V_i(\boldsymbol{r}-\boldsymbol{R}_j)]\psi_i(\boldsymbol{r}-\boldsymbol{R}_j)\, d\tau.$$

S'il y a n liaisons par maille, le déterminant a n lignes et n colonnes, donc l'équation (14) a n solutions en E : *les n liaisons par maille donnent naissance à n bandes d'énergie E.* S'il y a p atomes par maille, et comme il y a 1 état par maille dans chaque bande, on a ainsi en tout n/p états par atome, pouvant accepter *$2n/p$ électrons par atome.* Ainsi dans l'uranium α, $n = 4$, $p = 2$, d'où 4 électrons par atome dans la bande de liaison.

Le raisonnement précédent suppose les fonctions ψ_i *linéairement indépendantes.* C'est le cas si les fonctions de deux liaisons partant d'un même atome sont orthogonales. En première approximation, les fonctions des liaisons *asymétriques* du type de la figure 3 sont des combinaisons linéaires d'orbitales atomiques satisfaisant à cette condition.

Les fonctions ψ_i des liaisons *symétriques* du type de la figure 2 sont par contre, en première approximation, des combinaisons linéaires d'orbitales atomiques qui ne sont *pas* linéairement indépendantes; chaque orbitale atomique (hybride) sert pour au moins deux liaisons symétriques.

On sait que, dans ce cas, l'équation (18) est une identité en E. Pour avoir une équation qui ne soit pas identiquement satisfaite, il faut remplacer dans le développement (11) les orbitales de liaison ψ_i par un système $\Psi_{\boldsymbol{k}}$ de combinaisons linéairement indépendantes de ces fonctions. Le nombre de bandes $E(\boldsymbol{k})$, c'est-à-dire le nombre de lignes et de colonnes dans le déterminant de l'équation qui remplace alors (14), seront égaux au nombre par maille de fonctions $\Psi_{\boldsymbol{k}}$. Comme chaque fonction ψ_i utilise deux orbitales atomiques, le nombre des bandes se réduit ainsi à la moitié du nombre $m(< 2n)$ par maille d'orbitales atomiques linéairement indépendantes utilisées pour les liaisons. Ces bandes acceptent *m/p électrons par atome* s'il y a p atomes par maille. Dans l'uranium γ, $m = 3$, $p = 1$, d'où 3 électrons par atome dans la bande de liaison.

* Des mesures d'effet Hall permettraient de vérifier ce point.

VOL. 1 1956/57

RÉFÉRENCES

1. BATES L. F. et HUGHES D. *Proc. Phys. Soc.* **67,** 28 (1954).
2. GOODMAN B. B. et SHOENBERG D. *Nature* **165,** 441 (1950).
3. KATZ J. J. et RABINOVITCH E. *Chemistry of Uranium*, part I, New York (1951).
4. SOMMERFELD et BETH *Handb. der Physik* **24,** 2, 476 (1933).
5. STONER E. C. *Proc. Roy. Soc.* **A154,** 956 (1936).
6. VAN VLECK J. H. *Electric and Magnetic Susceptibilities*, Oxford (1932).
7. SEEGER A. Les électrons dans les métaux, 10ème congrès Solvay, Bruxelles (1955).
8. FRIEDEL J. *Proc. Phys. Soc.* **66A,** 333 (1953); à paraître (1957).
9. STONER E. C. *Acta Met.* **2,** 259 (1954).
10. WOHLFARTH E. P. *Phil. Mag.* **32,** 374 (1951).
11. KRIESSMAN C. J. et CALLEN H. B. *Phys. Rev.* **94,** 837 (1954).
12. MOTT N. F. et JONES H. *Metals and Alloys*, Oxford (1936).
13. THOMPSON J. G. *Nat. Bur. Stand. Rep. C.T.* 890, 16 (1943).
14. DAHL A. I. et VAN DUSEN M. S. *J. Res. Nat. Bur. Stand.* **39,** 53 (1947).
15. PAULING L. *J. Am. Soc.* **69,** 542 (1947).
16. DEHLINGER U. 1956 Conférence Stuttgart.
17. KIMBALL G. F. *J. Chem. Phys.* **3,** 560 (1935).
18. KIESS C. C., HUMPHREYS C. J., et LAUN D. D. *J. Res. Nat. Bur. Stand.* **37,** 57 (1946).
19. MOORE C. E. Atomic Energy Levels, Washington (1948).
20. KRUTTER H. M. *Phys. Rev.* **48,** 664 (1935).
21. SLATER J. C. et KOSTER G. F. *Phys. Rev.* **94,** 1498 (1954).
22. CAUCHOIS Y. *J. Phys.* **13,** 113 (1952).
23. GANZHORN K. Dissertation T. H. Stuttgart (1954).
24. *The Reactor Handbook* Vol. 3, Sect. I : Gen. prop. of mat. AECD 3647 (1955).
25. MACDONALD D. K. C. Bristol Conference, The Physical Society, Londres, p. 383 (1954).

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 188–190.

SUPERCONDUCTIVITY AND ELECTRON CONCENTRATION

B. T. MATTHIAS, E. A. WOOD, E. CORENZWIT, AND V. B. BALA

Bell Telephone Laboratories, Murray Hill, New Jersey, U.S.A.

(*Revised* 9 *August* 1956)

Abstract—Several systems are described in which the small additions of an element or compound lead to substantial changes of superconducting transition temperature. The conditions make it likely that the main effect is the change of electron concentration. In particular, molybdenum, reported to be a nonsuperconductor, is rendered superconducting by small percentages of rhodium.

VOL. 1 1956/57

THE purpose of this paper is to show how superconducting transition temperatures of certain systems can be varied by essentially changing the electron concentration, while the average mass and volume stay nearly constant.

The substances prepared in the course of this work were either melted in the helium arc furnace or reacted in sealed quartz tubes. The compositions given are those of the mixed elements before reaction.

(*a*) *Molybdenum*

This metal is reported to be normal down to 0·05°K.[1] The intention was to dissolve *small* amounts of another metal in molybdenum thus changing its electron concentration without actually affecting either its mass or volume, to the point at which molybdenum would become superconducting. One of the metals that could be dissolved and would change the electron concentration sufficiently was rhodium. From measurements of lattice constants RAUB[2] has concluded that rhodium is insoluble in molybdenum. However, the radii of the two elements are so similar if determined on the basis of co-ordination number eight that small amounts of rhodium could not have been detected in this way. The occurrence of superconductivity in molybdenum which was melted with small amounts of rhodium and then quenched may support our hypothesis and may demonstrate solubility of rhodium in molybdenum as well. Fig. 1 shows the variation of transition temperature with the percentage of rhodium dissolved in molybdenum. If the effect is really due to dissolved rhodium, then the solubility limit seems to be near 15 atomic per cent.

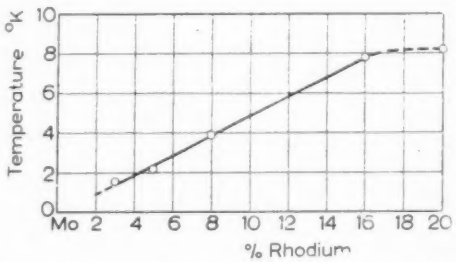


FIG. 1. Superconductivity of molybdenum-rhodium melts.

This is the first time to our knowledge that it has become possible to render a normal metal superconducting without any phase changes at all being involved.

(*b*) *The* β-W structure

It has been mentioned earlier[3] that the beta wolfram structure seems to be favorable for the occurrence of superconductivity. The discovery of several new compounds in this structure makes it possible now to trace the variation of the superconducting transition temperature, with the electron concentration as the only main variable, while mass and volume again remain practically unchanged. In Fig. 2 the compounds so related are

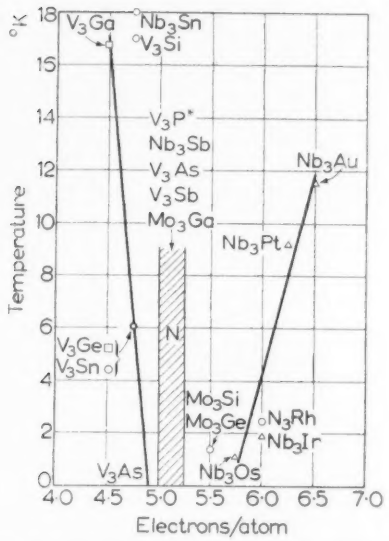


FIG. 2. Superconductivity in the β-W Structure. The V and Nb compounds connected are nearly identical in volume and mass. The o compounds are isolated as their corresponding neighbors fall in the nonsuperconducting range *N*.

* Slightly distorted.

connected by solid lines. Some of their lattice constants are listed by GELLER.[4] The lattice constants of the compounds discovered in the course of this work are given in Table 1.

Table 1

Substance	a_0 (Å)	Reference
Nb_3Rh	5·115	(8)
Nb_3Sb	5·26	
V_3Ga	4·83	
V_3Sb	4·92	
Mo_3Ga	4·94	
V_3As	4·74	(9)
V_3P	see reference [10]	(10)

(*c*) *The double carbides*

MoC and WC are hexagonal. However, they do form extended cubic solid solutions with the cubic carbides.[5] These cubic solutions, if quenched, cover in some instances a range up to 90 atomic per cent of the hexagonal constituent. Many of these solid solutions exhibited a range of lattice constants of as much as 0·1 Å from a value belonging to the greater part of the material to another representing the limit of substitution. From these solid solutions one might even extrapolate to the hypothetical sodium chloride modification of MoC. According to HARDY and HULM[6] the only superconducting carbides are those of niobium and molybdenum. In all others either electron concentration or too small a volume seem to lower their superconducting transition points below any accessible value. However, solid solutions between a number of pairs of carbides provide a larger range for the critical factors and therefore show superconductivity within a limited region.

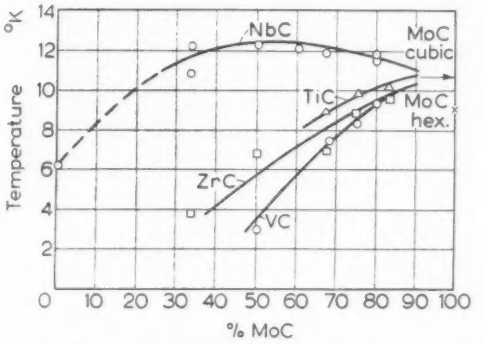


FIG. 3. Superconductivity of solid solutions between cubic carbides and molybdenum carbide.

Fig. 3 and Fig. 4 illustrate this behavior. Fig. 3 indicates that the transition temperature with increasing molybdenum content approaches gradually 10·6°K which is thus the transition temperature of the hypothetical cubic modification of MoC.

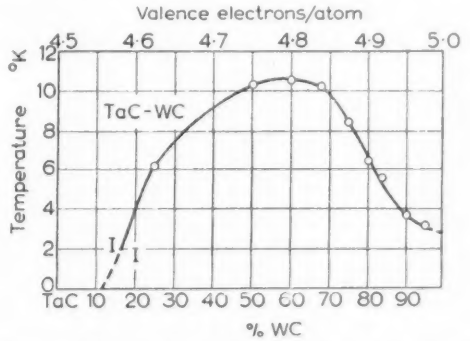


FIG. 4. Superconductivity in the TaC–WC system.

Until now only one superconductor has been known to show the phenomenon in two different crystal modifications. This is lanthanum in its cubic and hexagonal form.[7] The transition of the hexagonal form is about 1/2° higher than that of the cubic modification. As they are both close-packed structures, this is understandable. In the case of MoC the transition temperature of the hypothetical cubic modification is about 1·4° higher than that of the hexagonal one.

(d) Conclusion

In pseudosingle, binary, and ternary systems the superconducting transition temperature can be varied between wide limits by changes of the average valence electron concentration per atom. The crystal system, with the exception of the β-W structure which is particularly favorable, does not seem to be as important as the electron concentration.

We would like to express our gratitude to Drs. KIEFFER and BENESCHOVSKY for sending us TiC–WC samples which were also superconducting and thus stimulated the double-carbide investigation.

REFERENCES

1. SHOENBERG D. *Superconductivity* Cambridge University Press (1952).
2. RAUB E. *Z. Metallkunde* **45,** 23 (1954).
3. MATTHIAS B. T. *Phys. Rev.* **97,** 74 (1955).
4. GELLER S. A set of effective CN(12) radii for the β-wolfram-structure elements (1956) *Acta Cryst.* in press.
5. KIEFFER R. and SCHWARZKOPF P. *Hartstoffe & Hartmetalle* Springer Verlag (1953).
6. HARDY G. F. and HULM J. K. *Phys. Rev.* **93,** 1005 (1954)
7. JAMES N. R., LEGVOLD S., and SPEDDING F. H. *Phys. Rev.* **88,** 1092 (1952).
8. GREENFIELD P. and BECK P. A. *J. Metals* 265–276 (1956).
9. BACHMAYER K. and NOWOTNY H. *Monatsh. f. Ch.* **86,** 741 (1955).
10. BLITZ W. ET AL. *Z. anorg. Ch.* **249,** 21 (1942). (No lattice constant is given.)

VOL. 1 1956/5

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 191–199.

SELF-DIFFUSION AND ELECTRICAL CONDUCTIVITY IN SILVER CHLORIDE*

W. DALE COMPTON† AND ROBERT J. MAURER

Department of Physics, University of Illinois, Urbana, Illinois

(*Received* 26 *July* 1956)

Abstract—Measurement of the electrical conductivity and the self-diffusion coefficients of silver and chlorine were made in silver chloride. Self-diffusion coefficients were calculated from the conductivity by the Einstein relation. A discrepancy of 1·7 between the measured silver-diffusion coefficients and the calculated coefficients was found at 350°C. This discrepancy is interpreted as evidence that the migration of interstitial silver takes place by the *interstitialcy* mechanism. A discrepancy of 1·2 to 1·3 between the measured silver-diffusion coefficients and the calculated coefficients was found in the structure-sensitive region. This discrepancy is interpreted as evidence that silver-vacancy motion predominates in this temperature range. The measured diffusion coefficient of the chloride ion was found to be about 10^{-3} that of the silver ion.

INTRODUCTION

THE nature of the ionic defects present in the silver halides remained uncertain for many years.[1] Calculations by JOST and NEHLEP[2] and MOTT and LITTLETON[3] indicated that Frenkel defects are favored over Schottky defects but this was not definitely established by the calculation. Early experiments by TUBANDT[4] and more recent ones by KURNICK[5] established that the transport number of silver in silver chloride and silver bromide is unity for temperatures ranging from 20°C to above 400°C.

Measurements of electrical conductivity of pure silver halides were made by LEHFELDT.[6] Similar measurements on both pure crystals and crystals doped with positive divalent impurities were made by KOCH and WAGNER,[7] TELTOW,[8] and more recently by EBERT and TELTOW.[9] They found that the conductivity decreased and then rose nearly linearly as the concentration of divalent impurity was increased for temperatures where the crystal was initially intrinsic. This was explained by assuming that two types of mobile defects were present, and that the addition of the impurity first decreased the number of the more mobile carriers and increased the number of the less mobile carriers. By combining the above results with the mass-transport data, it was possible to give a very strong argument for the existence of Frenkel defects, where the interstitial silver ion has a higher mobility than does the silver ion vacancy.

MITCHELL[10] questioned this conclusion and suggested that Schottky defects exist in these crystals. He based his conclusions on an argument which sought to explain certain photographic phenomena and the growth of dendrites into a crystal from an electrode.

Several investigators then undertook experiments that were designed to give further information about the nature of the defects. STRELKOW's[11] thermal expansion data were analysed first by LAWSON[12] and later by CHRISTY and LAWSON.[13] ZIETEN[14] made new measurements of the thermal expansion of silver bromide. CHRISTY and LAWSON[13] and KANZAKI[15] made independent measurements of the high-temperature specific heats. The interpretation of the thermal-expansion data and the high-temperature specific-heat data requires a knowledge of the temperature dependence of the coefficient of

* Based on a Ph.D. thesis submitted to the University of Illinois. Partially supported by the National Science Foundation and the Office of Naval Research.

† Now at the Naval Research Laboratory, Washington, D.C.

thermal expansion and the specific heat of the normal crystal. Uncertainty in these quantities makes the conclusions rather uncertain. KURNICK[5] measured the change of conductivity of silver bromide as a function of pressure. He concluded that Frenkel defects exist below 300°C and a mixture of Frenkel and Schottky defects exist above 300°C. BERRY[16] measured the lattice parameter of silver bromide at room temperature and between 350 and 420°C. He found that the change in lattice parameter was sufficient to predict the external thermal expansion observed by STRELKOW[11] if no more than 10 per cent of the defects present were of the Schottky type. This was within the accuracy of the experiment. Measurements of the density defect in the silver halides also led BERRY[17] to conclude that the contribution of Schottky defects was negligible. FRIAUF[18] measured the polarization effects in the ionic conductivity of silver bromide and concluded that Frenkel defects were responsible for them.

Thus, the preponderance of experimental evidence supports the conclusion that Frenkel defects predominate in the silver halides.

Ionic motion can also be studied by measuring the self-diffusion of radioactive isotopes into the crystal. Diffusion measurements in pressed pellets of silver bromide at 300°C were reported by MURIN and TAUSH.[19] Such information was not available on single crystals, however. The following experiments were undertaken to gain this information by measuring electrical conductivity and self-diffusion of the constituents on single crystals having the same thermal history.[20]

THEORY

A crystal in thermal equilibrium contains lattice defects, in this case of the Frenkel type, which interact with the ions occupying normal lattice sites. Some of the lattice ion sites can also be occupied by impurity ions; for example, positive divalent impurities can be substituted for the cations. The crystal must remain electrically neutral, however. This is accomplished if a cation vacancy is introduced for each positive divalent impurity ion. If Y is the concentration of the divalent impurity, N_V the concentration of cation vacancies and N_i the concentration of cation interstitials then

$$N_V = N_i + Y. \tag{1}$$

A statistical mechanical argument[21] predicts the temperature dependence of N_V and N_i to be

$$N_V N_i = N_0^2 \exp(-W_0/kT), \tag{2}$$

where N_0 is a constant somewhat greater than the concentration of lattice sites, T is the absolute temperature, k is BOLTZMANN'S constant, and W_0 is the energy to form a defect pair at absolute zero. It is seen from (1) and (2) that temperatures may exist for which the number of defects generated thermally are so much greater than those introduced by the impurity that N_V and N_i are essentially equal. At these high temperatures the effect of the impurity is negligible and the crystal is considered to be pure. For sufficiently low temperatures, the number of defects generated thermally is negligible in comparison to those introduced by the impurity, so N_V and Y are almost equal.

The high-temperature range discussed above is called the intrinsic region and the low-temperature range is called the structure-sensitive region.

If a defect is to move from one lattice site to a neighboring site, it must surmount a barrier of height U. The defects execute a Brownian-type motion until an electric field is applied to the crystal. If (a) is the distance the defect moves in one jump, then the external field alters the barrier to $U \pm eEa/2$ such that a net flow of charge occurs in the direction of the field. For a field less than a few thousand volts/cm, the mobility of the defects can be derived from the net flow of charge in the field, and is given by

$$\mu = (\mu_0/T) \exp(-U_0/kT) \tag{3}$$

where μ_0 is a constant and U_0 is the height of the barrier at absolute zero which the defect must surmount in making a jump.

If each defect is of charge e, then the conductivity σ is given by

$$\sigma = e(N_V \mu_V + N_i \mu_i). \tag{4}$$

In the structure sensitive region N_i is negligible, N_V is constant, and the temperature variation of the conductivity is the same as the temperature variation of the mobility. In the intrinsic region N_i and N_V are equal but vary with temperature so the temperature variation of the conductivity is given by the product of the mobility and N_V or N_i. The slope of a $\ln \sigma T$ vs. $1/T$ curve is different

VOL. 1 1956/5

for the two regions. The temperature at which the slope changes is called the "knee" of the curve.

The motion of ions from one portion of a crystal to another can be described by a diffusion constant D. Since the transport occurs by a jumping of the defects, it is not surprising that the conductivity and diffusion are related. The Einstein relation[21] gives their dependence to be

$$\frac{\sigma}{D}=\frac{Ne^2}{kT} \tag{5}$$

where N is the concentration of lattice sites. Equation (5) implies that the motion of each defect contributes equally to the conductivity and diffusion. Thus, if the diffusion constant of only one constituent is to be compared with the conductivity, only that portion of the conductivity arising from this constituent should be used. The conductivity of a constituent is related to the total conductivity by the transport number n_t. BARDEEN and HERRING[22] have pointed out that (5) must also be modified when a measurement of D is made by a technique in which certain ions are tagged. After a tagged ion has jumped to a new location, one must consider whether its next jump is completely random. If the next jump has no correlation with the previous jump, then no correction to (5) is necessary. However, this may not be true and a factor f which is less than or equal to unity is introduced. Thus,

$$f\frac{n_t\sigma}{D_t}=\frac{Ne^2}{kT} \tag{6}$$

where D_t is the diffusion coefficient measured with radioactive tracers.

EXPERIMENTAL PROCEDURE

Crystals

Single crystals were grown in either air or helium by the Bridgmann method in 15-mm Vycor crucibles at a descent rate of 2·8 mm/hr. The starting material was either chemical reagent grade or Harshaw Chemical Co. silver chloride.

Sections were cut from the crystals, polished, and then annealed in air or helium for about six hours at 420°C. They were cooled to room temperature at a rate of 20°C/hour.

All crystals were handled in red light after growing.

Conductivity

The crystal was considered to have an equivalent circuit of a parallel resistance and capacitance. The resistance was measured with a General Radio 716-C capacitance bridge, used as an a.c. Wheatstone bridge, at frequencies between 10^2 and 10^4 cycles/sec. Since no frequency dependence was found, most measurements were taken at 10^3 cycles/sec. A helium atmosphere was maintained in the conductivity oven. Silver conducting paint was used for electrodes on the crystal. Temperatures were measured with a Pt and Pt+10 per cent Rh thermocouple.

Diffusion

It is assumed in a self-diffusion measurement utilizing radioactive tracers that the crystal cannot differentiate between the radioactive and nonradioactive ions. Thus, a concentration gradient of radioactive ions can exist without a concentration gradient of the constituents. Under these conditions the diffusion coefficient D appearing in FICK's laws is not a function of concentration. If C_0 is the initial surface concentration of radioactive ions, (t) the time taken for the diffusion and (h) the initial thickness of the radioactive layer deposited on a semi-infinite cylinder, then the solution of FICK's second law for $h \to 0$ is

$$C(X,t)=\frac{C_0}{\sqrt{(\pi Dt)}}\exp(-X^2/4Dt). \tag{7}$$

X is the distance from the radioactive surface layer. The boundary conditions are approximated by

$$h \ll \sqrt{(4Dt)} \ll L \tag{8}$$

for a crystal of thickness L.

From (7) it is evident that the value of D for a given temperature and time of diffusion is obtained from the slope of a plot of the logarithm of the number of radioactive ions in a unit thickness vs. X^2.

The Ag^{110} isotope of half-life of about 270 days was placed on the surface of the crystals by evaporation or by chemically depositing it from solution. In the former method the radioactive silver chloride was prepared from radioactive silver nitrate. In the latter method the layer of silver containing the Ag^{110} was converted to silver chloride by fresh chlorine water. The radioactive silver chloride containing the Cl^{36} isotope of half life of approximately 4×10^5 years was prepared from radioactive hydrochloric acid and was evaporated onto the surface of the crystals.

The diffusions were carried out in a helium atmosphere. Temperatures were measured with Pt and Pt+10 per cent Rh thermocouples. Corrections were applied for the heating and cooling portions of the diffusion cycle. Slices were taken from the crystals with a microtome by the procedure outlined by MAPOTHER, CROOKS, and MAURER.[23] Fig. 1 shows a typical profile plot of the activity vs. X^2. The penetration was usually about 100 micron, the thickness of the radioactive layer was about 0·2 micron, and the crystal length was about 1 cm. Thus, the conditions of (8) were satisfied experimentally. Penetration was less for the chlorine measurements since lower initial activities were used.

Temperature

The Pt and Pt+10 per cent Rh thermocouples were calibrated with National Bureau of Standards freezing-point samples of lead, tin, and zinc. For temperatures below 100°C calibration was carried out by comparison with a new mercury thermometer certified by the National Bureau of Standards.

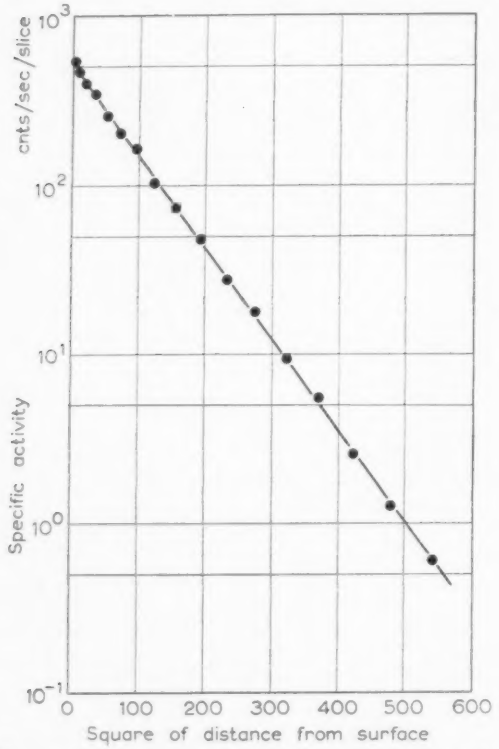


FIG. 1. Specific activity vs. distance squared from the surface.

Distance below the surface of the crystal measured in units of $4{\cdot}0\times10^{-4}$ cm. The activity is plotted at the center of the slice. Diffusion temperature = 268·0°C. Diffusion time = 1021 seconds. $D = 3{\cdot}2\times10^{-9}$cm²/sec.

Accuracy

The accuracy of the conductivity measurements depends upon the measurement of temperature, crystal dimensions, and resistance. The total combined errors are considered to be less than ±10 per cent. The diffusion constant results are subject to errors in the measurement of temperature, diffusion time, determination of the slope of the ln activity vs. X^2 curve, and misalignment of the sample when the slices are taken. A total combined error of less than ±10 per cent was maintained for the silver diffusion and about ±15 per cent for the chlorine diffusion measurements.

RESULTS

Results are shown for four crystals in Figs. 2, 3, and 4. Crystals 18b and 22a were single crystals while crystals 19a and 19b contained two and four subcrystals, respectively. In all cases the diffusion coefficients measured with the silver radioisotopes were smaller than coefficients calculated from (6) with n_t and f taken as unity. At 350°C this discrepancy was about 1·7 for all crystals, while at 200°C it was about 2·6 for crystal 18b and 2·1 for the others. Intrinsic measured diffusion coefficients shown in Fig. 4 were found to agree for all four crystals. Thus, the intrinsic conductivity of crystal 18b was about 20 per cent greater than that of the other crystals.

Crystals 18 and 19 were grown and annealed in air from Harshaw material. Crystal 22 was grown from Baker reagent grade material in helium and annealed in air. It was found that the conductivity of Harshaw crystals and of crystals grown and annealed in helium agreed with the conductivity of crystals 19 and 22. These results are in agreement with results published by EBERT and TELTOW.[9] Crystal 18b had an anomalously high conductivity.

Chlorine diffusion measurements were made on the same crystals. The results appear on Fig. 4.

Fig. 5 gives the results of calculated and measured diffusion coefficients on a crystal doped with 0·015 mole per cent of cadmium chloride. The numbered points indicate the order in which the diffusion measurements were carried out and indicate that an annealing effect took place after points 1 and 2 were taken. The measured diffusion coefficient is about 30 per cent less than the calculated coefficient in the structure sensitive region.

A least-square analysis of the data gives the activation energies listed in Table 1.

Table 1. *Activation Energies*

	Measured with Radioisotopes	
0·890 eV	Silver diffusion	Fig. 4
1·61 eV	Chlorine diffusion	Fig. 4
	Calculated from Conductivity	
0·884 eV	Intrinsic silver diffusion	Fig. 2
0·835 eV	Intrinsic silver diffusion	Fig. 3
0·38 eV	Structure-sensitive silver diffusion	Fig. 2
0·36 eV	Structure-sensitive silver diffusion	Fig. 5

VOL. 1 956/5

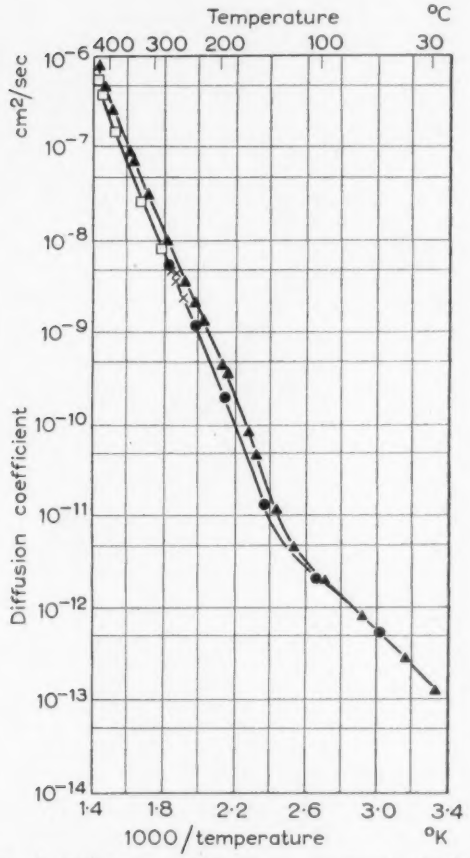


FIG. 2. Diffusion coefficients vs. reciprocal absolute temperature in silver chloride. ▲—Diffusion coefficients calculated from the electrical conductivity by (6) with $f = 1$. ◇, ●, ×—Diffusion coefficients of silver measured with Ag^{110} radioisotope. ◇—crystal 19b. ●—crystal 19a. ×—crystal 22a.

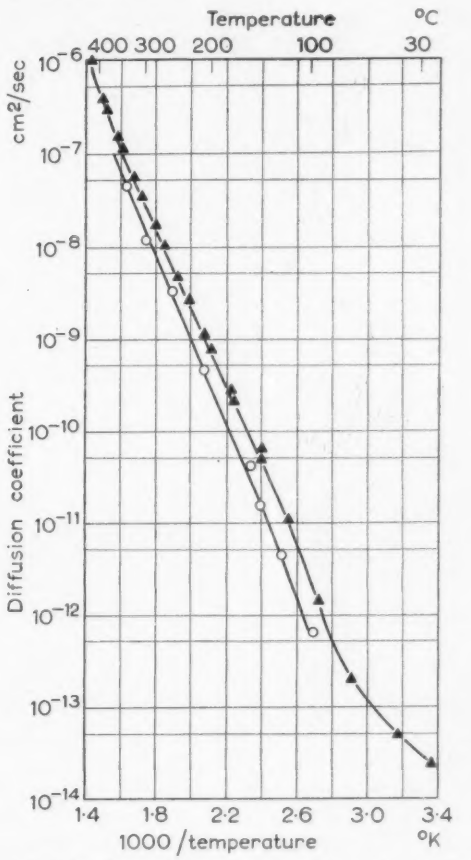


FIG. 3. Diffusion coefficients vs. reciprocal absolute temperature in silver chloride. ▲—Diffusion coefficients calculated from the electrical conductivity by (6) with $f = 1$. O—Diffusion coefficients of silver measured with Ag^{110} radioisotope-crystal 18b.

DISCUSSION OF RESULTS

Results of previously mentioned experiments[4, 5] and of the present chlorine diffusion measurement allow one to set n_i equal to unity for silver in (6) and zero for chlorine and electronic transport. Since silver transport takes place by vacancies and interstitials, it is necessary to consider whether f is the same for both types of transport.

Theoretical calculations by HOVE[24] indicate that the energy required for a silver interstitial to migrate through a cell face is much greater than that required for it to displace a neighboring ion to an interstitial site and then occupy that vacant lattice site. The latter type of motion has been called the *interstitialcy* mechanism by KOCH and WAGNER[25] and SEITZ.[26] MCCOMBIE and LIDIARD[27] have calculated f for the *interstitialcy* type of motion. They point out that for this type motion charge is transported a distance $2d$ while a radioactive ion only moves a distance d. Thus, a factor of 0·5 should appear on the left side of (6) for the *interstitialcy* mechanism. This factor will be included in f although it does not arise from correlation effects. These authors also point out that an ion that has just jumped from an interstitial site to a lattice site will have a high probability of

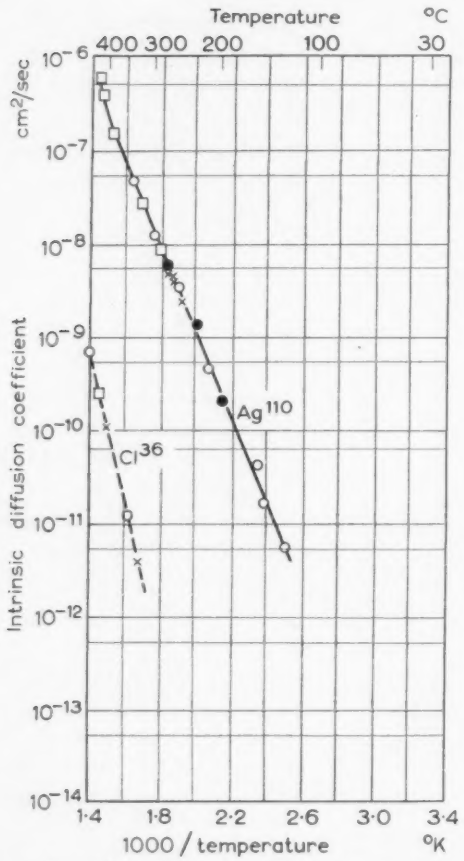


FIG. 4. Instrinsic diffusion coefficients vs. reciprocal absolute temperature in silver chloride. —— Diffusion coefficients of silver measured with Ag^{110} radioisotope on crystals 18b, 19a, 19b, and 22a. - - - - Diffusion coefficients of chlorine measured with Cl^{36} radioisotope on crystals 18b, 19b, and 22a.

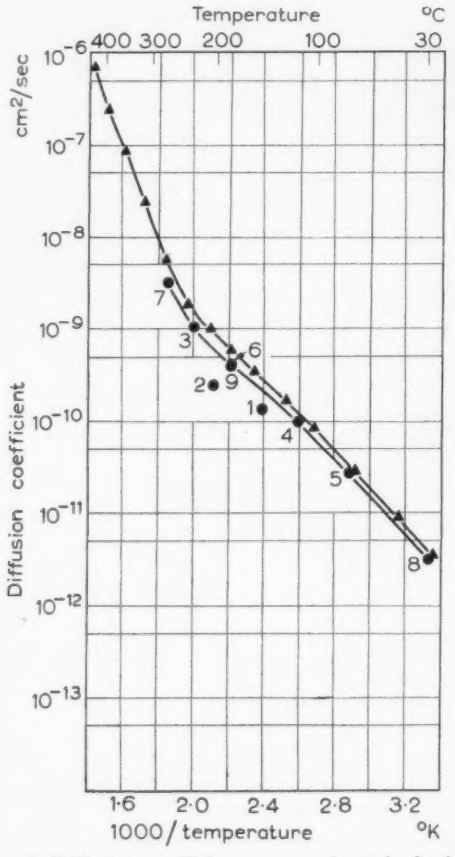


FIG. 5. Diffusion coefficients vs. reciprocal absolute temperature in silver chloride doped with 0·015 mole per cent $CdCl_2$. ▲—Diffusion coefficients calculated from the electrical conductivity by (6) with $f = 1$. ●—Diffusion coefficients of silver measured with Ag^{110} radioisotope.

returning to the original interstitial site because of the location of the interstitial next to it. On the other hand, the jump of an ion on an interstitial site should not be correlated with its previous jump. A calculation of these correlation effects, assuming that the lattice site and the two interstitial positions are collinear, gives a factor 0·67 and the combined result for the *interstitialcy* mechanism is

$$f_i = 0{\cdot}33.$$

BARDEEN and HERRING[22] found for vacancy motion that

$$f_V = 0{\cdot}80.$$

Thus, from (6)

$$D_i = (kT/Ne)(f_i N_i \mu_i + 0{\cdot}80 N_V \mu_V). \qquad (9)$$

Neglecting the point for the lowest temperature, the data of EBERT and TELTOW[9] can be represented by

$$\mu_V/\mu_i = 41{\cdot}6 \exp(-0{\cdot}272/kT). \qquad (10)$$

Substituting this into (9) and (4) and considering the intrinsic region where $N_i = N_V$ gives

$$D_i = (kT/Ne)N_i\mu_i\{f_i + 33{\cdot}3 \exp(-0{\cdot}272/kT)\} \qquad (11)$$

and

$$\sigma = N_i e \mu_i \{1 + 41{\cdot}6 \exp(-0{\cdot}272/kT)\}. \qquad (12)$$

The last term of (12) varies from 0·053 at 200°C to 0·26 at 350°C. This is a slow variation with temperature and a plot of $\ln \sigma T$ vs. $1/T$ for the intrinsic region between these two temperatures should still give a straight line with an activation energy determined by $N_i\mu_i$. This is the significance of the intrinsic silver activation energies of Table 1.

The lower curves of Figs. 2 and 3 are diffusion coefficients measured with radioactive tracers. The upper curves of these figures are diffusion coefficients calculated from the conductivity by (6) with f and n_t set equal to unity. Since f is not unity, the magnitude and the temperature dependence of the discrepancy between the two curves should be predictable from the ratio of (6) and (11) if (12) is substituted for σ in (6). This ratio is given by

$$R = \frac{1+41{\cdot}6\exp(-0{\cdot}272/kT)}{f_i+33{\cdot}3\exp(-0{\cdot}272/kT)}. \qquad (13)$$

The data of Fig. 2 are best represented by taking $f_i = 0{\cdot}50$ rather than 0·33 as McCombie and Lidiard[27] predicted. This gives $R = 1{\cdot}78$ at 350°C and $R = 1{\cdot}94$ at 200°C. The observed R varies from 1·7 at 350°C to 2·1 at 200°C. The observed R in Fig. 3 varies from 1·7 at 350°C to 2·7 at 200°C. The difference in R at 200°C between Figs. 2 and 3 results from the different slopes as calculated from the conductivity and as given in Table 1. This difference is not understood.

McCombie and Lidiard[27] have pointed out that the discrepancy between their calculated value of f_i and the value required to give agreement with experiment is not surprising since they limited their calculations to a collinear jump mechanism. However, Hove's[24] calculations indicate that the energy for the collinear jump is much less than for a noncollinear jump. In fact, Hove's results imply that the stable configuration is not an interstitial and an ion on a normal lattice site, but consists of the two silver ions symmetrically spaced about the plane containing the normal lattice. It is not unlikely, however, that the movement of a unit such as this would have a correlation factor f_i which is smaller than 0·33 rather than the experimentally observed larger value.

McCombie and Lidiard[27] emphasize that the observed discrepancy cannot be the result of a local field effect which tends to give an anomalously high conductivity. This was suggested by Katz.[28]

In the structure sensitive region N_i is negligible and $N_V = Y$ so

$$D_t = 0{\cdot}80(kT/Ne)Y\mu_V. \qquad (14)$$

The low temperature data of Figs. 2 and 5 were calculated with the factor of 0·80 in (14) replaced by unity. The resulting discrepancy between the measured and calculated diffusion coefficients is between 1·2 and 1·3. Thus, these data support the conclusion that vacancy motion predominates in the structure sensitive region. The agreement of the two curves at 58°C in Fig. 2 is believed to result from accidental contamination of the crystal during previous measurements.

A straight line is evident in the intrinsic region of Fig. 2 between 150°C and 350°C. An extension of this straight line to lower temperatures shows that the diffusion coefficients drop below this line prior to the temperature at which the curve has the slope characteristic of the structure sensitive region.

The drop at low temperatures results from the removal of the more mobile interstitials by positive divalent impurities. A decrease in the temperature in the region just above the knee of the conductivity curve increases the influence of the impurity and tends to remove interstitials as is predicted from (1) and (2). The effect is substantially equivalent to that observed in the experiment of Ebert and Teltow,[9] in which the conductivity initially decreased as the concentration of divalent impurity was increased for a constant temperature.

An extension of the straight line of the intrinsic region to temperatures above 350°C emphasizes that the silver diffusion and conductivity rise anomalously, but that they remain approximately parallel. This rise is not understood but a few observations are possible from the above data.

The contribution of the second term of (11) and (12) becomes important at these temperatures. However, calculations show that these terms can not produce the abrupt rise observed in the data over this small temperature range if the activation energies for the motion of the defects remain unchanged.

It has been suggested that the generation of Schottky defects could produce such a rise. If the chlorine defects are mobile, then the measured chlorine diffusion implies that they contribute

negligibly to the total conductivity. Application of statistical mechanics shows that as the concentration of chlorine vacancies increases the concentration of silver interstitials must decrease. This would result in an initial drop in the conductivity followed by a subsequent rise. Such a decrease is not observed.

As EBERT and TELTOW[9] emphasize, the most important influence may be the thermal expansion of the lattice. Since BERRY[16] has shown that the volume thermal expansion results from an increase in lattice parameter, it seems necessary to consider the effect that this has upon the pre-exponential constants in the equations for conductivity and diffusion.[21] Data in this temperature range are not sufficient to allow more reliable conclusions to be drawn.

SUMMARY OF RESULTS

Self-diffusion measurements of silver and chlorine into four crystals of silver chloride were made with Ag^{110} and Cl^{36} radioisotopes. For the intrinsic region below 350°C, the silver-diffusion data are represented by

$$D_{Ag} = 1{\cdot}46 \exp(-0{\cdot}890/kT)\ \text{cm}^2/\text{sec}.$$

The chlorine diffusion data for temperatures between 324°C and 443°C are represented by

$$D_{Cl} = 133 \exp(-1{\cdot}61/kT)\ \text{cm}^2/\text{sec}.$$

Electrical-conductivity measurements were made on the same crystals. The activation energies for the intrinsic region were found to be 0·884 ev for all crystals except one. The activation energy for motion of the silver vacancy was found to be 0·36 ev. These values are in substantial agreement with those obtained by other investigators.[7, 9]

Diffusion coefficients calculated from the electrical conductivity by the Einstein relation are larger than the measured diffusion constants by 1·7 at 350°C[29] and 1·25 in the structure sensitive region.

These results are in good agreement with the assumption that the transport of silver occurs by the vacancy mechanism in the structure sensitive region and they suggest that, at high temperature, the silver ion moves by the "interstialcy mechanism." The data is not in quantitative agreement, however, with the predictions of the "interstialcy mechanism."

Acknowledgement—The authors wish to acknowledge the many stimulating discussions held with Dr. C. W. MCCOMBIE. One of the authors (COMPTON) wishes to express his gratitude to the U.S. Naval Ordnance Test Station, Inyokern, for the Station Fellowship which he held while working on a portion of this problem.

REFERENCES

1. SEITZ F.Speculations on the properties of the silver halides *Rev. Modern Phys.* **23,** 328 (1951).
2. JOST W. and NEHLEP G. *Z. physik. Chem.* **32B,** 1 (1936).
3. MOTT N. F. and LITTLETON M. J. *Trans. Faraday Soc.* **34,** 485, 500 (1938).
4. TUBANDT, C. *Z. anorg. u. allgem. Chem.* **115,** 113 (1920).
5. KURNICK S. W. *J. Chem. Phys.* **20,** 218 (1952).
6. LEHFELDT W. *Z. Physik* **85,** 717 (1933).
7. KOCH E. and WAGNER C. *Z. physik. Chem.* **38B,** 295 (1937–38).
8. TELTOW J. *Ann. Physik.* **5,** 63, 71 (1949).
9. EBERT I. and TELTOW J. *Ann. Physik* **15,** 268 (1955).
10. MITCHELL J. W. *Phil. Mag.* **40,** Series 7, 667 (1949).
11. STRELKOW P. G. *Physik. Zeits Sowjetunion* **12,** 77 (1937).
12. LAWSON A. W. *Phys. Rev.* **78,** 185 (1950).
13. CHRISTY R. W. and LAWSON A. W. *J. Chem. Phys.* **19,** 517 (1951).
14. ZIETEN W. *Z. Physik* **145,** 125 (1956).
15. KANZAKI H. *Phys. Rev.* **81,** 884 (1951).
16. BERRY C. *Phys. Rev.* **82,** 422 (1951).
17. BERRY C. *Phys. Rev.* **97,** 676 (1955).
18. FRIAUF R. *J. Chem. Phys.* **22,** 1829 (1954).
19. MURIN A. and TAUSH YU. *Doklady Akad. Nauk S.S.S.R.* **80,** 579 (1951).
20. COMPTON W. D. *Bull. Amer. Phys. Soc.* **30,** No. 2, 20 (1955) and *Phys. Rev.* **101,** 1209 (1956).
21. See for example MOTT N. F. and GURNEY R. W. *Electronic Processes in Ionic Crystals* (Oxford University Press, London, 1950), Chapter II.
22. BARDEEN J. and HERRING C. in *Imperfections in Nearly Perfect Crystals*, Edited, W. Shockley, Chairman. (John Wiley and Sons, Inc., New York, 1952), p. 261.
23. MAPOTHER, D., CROOKS H. N., and MAURER R. *J. Chem. Phys.* **18,** 1231 (1950).
24. HOVE J. E. *Phys. Rev.* **102,** 915 (1956).
25. KOCH E. and WAGNER C. *Z. physik. Chem.* **B38,** 295 (1937).
26. SEITZ F. *Acta Cryst.* **3,** 355 (1950).
27. MCCOMBIE C. W. and LIDIARD A. B. *Phys. Rev.* **101,** 1210 (1956).
28. KATZ E. *Phys. Rev.* **99,** 1334 (1955).
29. Similar results have recently been reported on silver bromide, FRIAUF R. J. *Bull. Amer. Phys. Soc.*, Ser. II, **1,** 145 (1956).

VOL. 1 1956/57

J. Phys. Chem. Solids. Pergamon Press 1956. Vol. 1. pp. 199–208.

EFFECT OF PRESSURE ON THE SUPERCONDUCTING TRANSITION OF LEAD*†

R. R. HAKE AND D. E. MAPOTHER

Physics Department, University of Illinois, Urbana, Illinois

(*Received* 16 *July* 1956)

Abstract—The effect of hydrostatic pressures up to 300 atmospheres on the critical magnetic field of superconducting lead has been measured at a temperature slightly below the critical temperature. The pressures were applied using helium gas as the pressure transmitting fluid. The critical fields of the pressurized specimen and a zero-pressure comparison specimen were measured alternately using a sensitive d.c. ballistic method. The pressure-induced critical field difference ΔH_c was found to vary linearly with pressure with a slope $(\partial H_c/\partial P)_T = -9{\cdot}60 \pm 0{\cdot}11 \times 10^{-9}$ gauss/dyne cm^{-2} at about 7·09°K. The factors influencing the magnitude of the pressure shift are discussed on the basis of the Bardeen theory. The pressure shift for lead, unlike that for other superconductors which have been measured, is about the same as that which is estimated from a consideration of zero-point vibration-amplitude changes only.

1. INTRODUCTION

THE isotope effect,[1, 2] discovered in 1950, showed that the superconducting transition temperature T_c of different isotopes of mercury varied approximately as the inverse square root of the average isotopic mass M. As was shown later, this is also the case for isotopes of tin and thallium.[3] It may be shown that the mean square amplitude of the zero-point lattice vibrations, $\langle q_0^2 \rangle_{Av}$, is also proportional to $M^{-\frac{1}{2}}$. This implies that the interaction which is responsible for superconductivity is between electrons and phonons. Moreover, since its strength, as measured by the magnitude of T_c, is proportional to $\langle q_0^2 \rangle_{Av}$, the interaction must be similar to that which is responsible for the high-temperature resistivity. Recent theories by BARDEEN[4] and FROHLICH[5] have sought to formulate this interaction, somewhat along the lines of ordinary conductivity theory. Various mathematical and physical difficulties appear, however, and no clear picture of the mechanism responsible for superconductivity has as yet been achieved.

Aside from the isotope effect, another experimentally direct method of investigating the interaction is to study the volume dependence of the superconducting transition.[6] Changes in volume, unlike changes in isotopic mass, might be expected to alter significantly not only the amplitude of the lattice vibrations but also the electronic character of a metal. The lowering of the transition temperature caused by compression of Sn, In, and Al[7] appears to be at least twice as large as would be expected if one assumes that T_c is proportional to $\langle q_0^2 \rangle_{Av}$ and estimates the decrease in lattice vibrational amplitude under compression from GRÜNEISEN'S constant (see eq. 18, sect. 4). For the superconductors Bi_4Rh, Bi_3Ni, and Bi_2K,[8, 9] on the other hand, the transition temperature actually increases under compression. For Tl[10] there is an increase in the transition temperature up to about 1500 atmospheres, while further increase in pressure reduces the transition temperature. The effect of pressure on the transition temperature is, then, evidently more dependent on the detailed nature of the electron-phonon interaction than is the effect of isotopic-mass variation. While this circumstance makes pressure-effect data of

* This work has received partial support from the Office of Ordnance Research, U.S. Army.

† This paper is based upon a thesis presented to the University of Illinois by R. R. HAKE in partial fulfillment of the requirements for the Ph.D. degree.

considerable potential value in indicating the important mechanisms affecting the superconducting interaction, for the present, interpretation is hindered by the lack of knowledge as to the effects of compression on the electronic energy-band structure and on the vibrational spectrum.

In this experiment we have measured the effect of hydrostatic pressure on the critical field of lead. Lead was chosen for this study because it was thought that its cubic structure might better facilitate calculation to determine the nature of the vibrational and electronic changes under compression. Its high transition temperature also makes possible the eventual measurement of pressure-induced shifts (and isotopic shifts) over a very wide temperature range, thus allowing a separation of effects due to electronic and vibrational factors as will be discussed more fully below.

2. APPARATUS AND EXPERIMENTAL METHOD

(a) *Method*

In order to measure the pressure-induced shift of the critical field of lead a comparison method was used. Two identical pure single crystals of lead were placed in a cryostat providing close thermal contact between the specimens at a temperature of about 7°K. Hydrostatic pressure was applied to only one specimen, the other specimen serving as a zero-pressure comparison. The pressure was transmitted through gaseous helium so that it could be applied and varied at low temperature and could be accurately measured by bourdon gauges in the room temperature portions of the apparatus. Measurements of the critical fields of the two specimens were made alternately, while the temperature and pressure were maintained constant. The difference in the critical fields, then, was the shift due to pressure. The measurement of the shift was repeated at different pressures so that the critical field shift was obtained as a function of pressure, up to the solidification pressure of helium at 7°K, about 300 atmospheres. The critical fields were determined using a sensitive d.c. ballistic method which can resolve critical fields to within a few hundredths of a gauss. The comparison method allowed easy detection of, and correction for, any temperature drift during the course of the measurement.

(b) *Cryostat*

The cryostat and its connections are shown schematically in Fig. 1. The specimens, carbon resister thermometer, constantan heater, and search coils were all located inside the "inner can", and affixed to a central copper tube so as to promote thermal contact with one another.

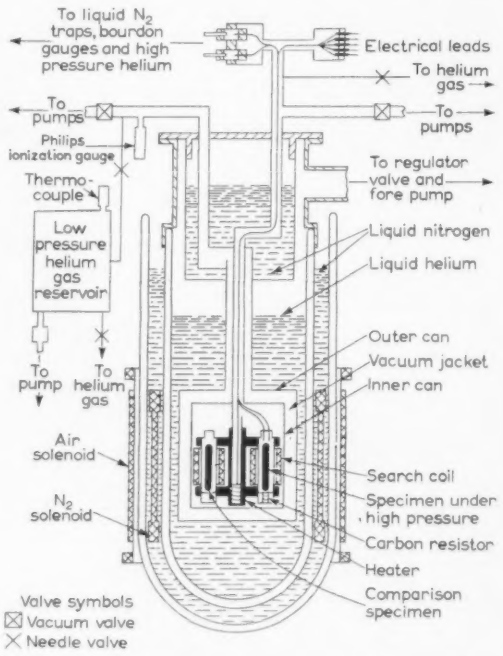


FIG. 1. Schematic diagram of cryostat and connections. The empty bomb and search coil are not shown.

The specimens were encased in beryllium-copper bombs. The bomb containing the comparison specimen was sealed off in an atmosphere of helium gas at room temperature, while the bomb containing the pressurized specimen was connected to the source of high-pressure helium gas. There were actually three identical bombs and search coils clamped symmetrically around the central copper tube. The third bomb and search coil, which are not shown in Fig. 1, were empty, serving only as a comparison dummy for the magnetic measurement.

Helium gas was admitted to the inner can in order to make it as nearly an isothermal enclosure as possible. The "outer can", which surrounded the inner can, was immersed in a bath of liquid helium. An outer dewar containing liquid nitrogen surrounded the liquid-helium dewar. The helium bath was pumped through a special bellows-actuated regulator valve so that its temperature, and therefore the temperature of the outer can, remained constant at about 4·0°K. The degree of thermal contact between the inner and outer cans could be adjusted by

VOL. 1 1956/57

controlling the exchange gas pressure in the vacuum jacket between the two cans. Fine control of the pressure in the vacuum jacket was made possible by the parallel arrangement of bleeding and pumping systems shown in Fig. 1.

The inner and outer cans could be opened via demountable joints which were sealed with a nonsuperconducting soft solder. The high-pressure tubes and electrical leads were brought out from the inner can through the thin-walled inconel tube which suspended the inner can from the nitrogen trap.

(c) *Critical field measurement*

With the relatively modest pressures available in the gaseous helium range below the solidus curve, the critical field shifts were only of the order of a gauss or so. Thus, in order to determine critical field shifts to within a few per cent, it was necessary to resolve critical fields to within a few hundredths of a gauss. The technique and apparatus with which such resolution was obtained were developed in this laboratory for critical field measurements on Al and are discussed fully by COCHRAN, MAPOTHER, and MOULD in a recent publication.[11] Only the briefest description of the method and apparatus will be given here.

The critical fields were determined by measuring the flux changes in the specimens as an external magnetic field was varied in small steps from a point in the superconducting state to a point in the normal state. The external field was generated by the two solenoids[11] shown in Fig. 1, each of which was specially designed to produce a field homogeneity better than 0·1 per cent over the region occupied by the specimens. The "N_2 solenoid" produced a steady magnetic field slightly less than the critical field, while the "air solenoid" produced a small additional field which could be changed in steps of about 0·08 gauss so as to bracket the transition. An electronic current regulator plus occasional manual adjustment kept the steady N_2-solenoid field constant to within 0·005 gauss at about 30 gauss, where the transitions were taken. The flux changes in the specimen in response to each incremental field increase were picked up by the search coils surrounding the three identical bombs (see Fig. 1). Each search coil consisted of 9751 turns of No. 40 Formex insulated copper magnet wire. Each of the two specimen search coils could be connected in series opposition to the dummy search coil, surrounding the empty bomb, through a high-sensitivity ballistic galvanometer. The galvanometer deflections were used to calculate an effective permeability μ_e,[11] which varies from 0 to 1 as the superconductor passes from the superconducting to the normal state. Fig. 2 shows μ_e plotted against the external applied field H_e for a superconducting to normal transition. For an ideal ellipsoidal specimen of the same major and minor diameter as the real specimen (the ratio of these diameters specifying the demagnetizing factor $n = 0{\cdot}0167$) μ (defined as B/H_c[11]) should vary in the manner shown by the dashed line.

It appears that the actual transition width, ignoring the rounding at the extremes of the transition, was very close to the ideal width, and was thus determined mainly by the geometry of the specimen. It is believed that the rounding at the extremes of the transition is due to secondary effects[11] and that the best experimental approximation to the true thermodynamic value of H_c is obtained by extrapolating the central straight-line portion of the μ_e vs. H_e curve to $\mu_e = 1$ as is shown in Fig. 2.

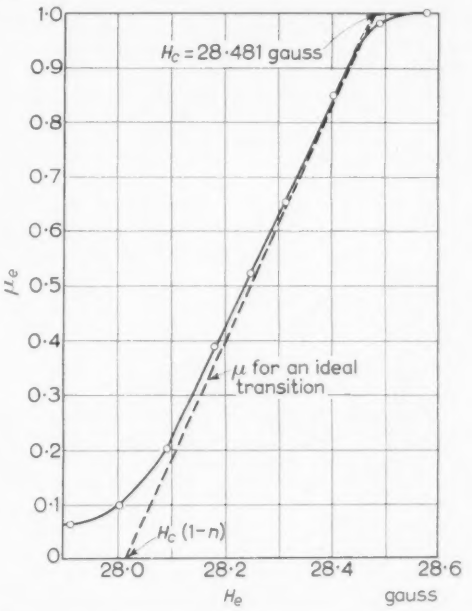


FIG. 2. A superconducting-to-normal transition for a lead specimen at constant temperature and pressure. The critical field H_c is obtained by extrapolation as shown.

(d) *Temperature control*

The steepness of the H_c vs. T curve for lead at 7°K is such that temperature changes of 10^{-4}K° can cause critical field shifts of about 0·02 gauss. Thus, in order to measure a critical field to within a few hundredths of a gauss, it was necessary to maintain the temperature constant to within a ten thousandth of a degree during the course of a critical field measurement lasting about ten minutes. This temperature constancy was obtained by careful control of exchange gas pressure in the vacuum jacket between the inner and outer cans, and by a carbon-resistor-actuated heater in the inner can (see Fig. 1).

By suitable manipulation of the exchange gas pressure and a steady heating current to the inner can, heat flow conditions could be established such that the inner can was at the desired temperature of about 7°K with a slight tendency to cool. The cooling tendency was balanced against a small intermittent regulating current to the heater. This intermittent heater current was actuated by the carbon-resistor thermometer so that it was turned

on and off as the temperature of the carbon resistor fell below or rose above the desired temperature. In this manner the temperature of the carbon resistor could be kept constant to within 10^{-4}K° at about 7°K.

The carbon-resistor thermometer was a 56-ohm Allen-Bradley radio resistor cemented into the bottom of the bomb containing the pressurized specimen as shown in Fig. 1. This resistor was made one arm of a standard Wheatstone bridge, which was set in balance when the carbon resistor was at the desired temperature. The null detector was a Leeds and Northrup 9835 B stabilized d.c. microvolt amplifier, which was used as a preamplifier for a type-G Leeds and Northrup Speedomax recorder. The sensitivity of the carbon resistor thermometer at 7°K was about $7{\cdot}3\times10^{-3}$K°/ohm. The amplifier gain was such that the temperature sensitivity in terms of movement of the recorder needle was about 10^{-4}K°/inch. When the recorder needle moved off balance on the cold side, a microswitch was thereby closed, sending about 0·01 mA into the heater (0·35 μW). This warmed the carbon resistor back to the balance point where the microswitch opened and the cycle was then repeated.

During periods of regulation the specimens themselves usually exhibited a slow and rather steady temperature drift. This drift was detected by the steady displacements of the critical fields in a series of successive measurements. The temperature drift appeared to be the same in the two specimens. It was usually possible to select for comparison from a series of alternate measurements, three successively measured critical fields such that the critical-field shift for the two measurements on the same specimen $\Delta H(D)$ was less than 0·03 gauss (see Table 1). This corresponds to a temperature drift of about $1{\cdot}5\times10^{-4}$K° during the three measurement comparison (about 30 minutes), and a drift of 5×10^{-5}K° during one critical-field measurement.

(e) *High-pressure technique*

Helium gas was compressed to about 400 atmospheres by means of a hand-operated gas booster pump ("Superpressure" Model 406-921, American Instrument Co.). The helium became contaminated with oil in the compression cylinder but this was effectively removed by two liquid-nitrogen traps in the high pressure input line leading to the cryostat. There were two separate high-pressure lines. The input line connected to the high-pressure pump through a series of valves which enabled the pressure to be set at any desired value. The output line connected to a series of large bourdon gauges on which the pressure was read. These were calibrated against a dead-weight tester to an accuracy of $\pm\frac{1}{2}$ psi. The calibrations taken before and after the run were identical. Two pressure lines were used so that blocks in the lines due to freezing of helium or of foreign matter would be clearly evidenced by a failure of the pressure gauges on the two lines to follow one another when the pressure was changed.

As Fig. 1 shows, the high-pressure tubes were brought down the central pumping tube and were thus well insulated from the surrounding liquid-helium bath. This is important since high-pressure helium would have solidified in the tubes if they had been anywhere near the 4°K bath temperature. The high-pressure tubes were *k*-monel and had an o.d. of 0·016 inch, and an i.d. of 0·008 inch. Such small diameter tubes were used in order to reduce heat influx.

The high-pressure bomb was machined from beryllium copper (Berlyco-25, $\frac{1}{2}$ hard). After machining, the bomb was heat-treated to increase its tensile strength to around 200,000 psi. The high-pressure tubes were soft soldered into the top of the high-pressure bomb. The bomb was sealed at the bottom by a teflon gasket joint, not shown in Fig. 1.

(f) *Specimens*

High-purity single crystals of lead were used in this study. They were in the form of rods with rounded ends. The high-pressure specimen was 5·60 cm long and had an average diameter of 0·493 cm, uniform to about 0·7 per cent over its length. The comparison specimen was 5·50 cm long and had an average diameter of 0·490 cm, uniform to about 0·7 per cent over its length. The lead was 99·999+ per cent pure and was obtained from the American Smelting and Refining Co. The crystals were grown under vacuum in graphite-coated glass tubing situated in a temperature-gradient furnace. Solidification was caused to proceed up the melt by slowly reducing the current applied to the furnace. The total time for the crystals to cool from the melting point to room temperature was about 10 hr. The crystals were removed from the glass tubing by dissolving the glass in hydrofluoric acid. The ends of the crystals were rounded to approximate ellipsoidal shape by controlled electrolytic dissolution. The crystals were polished by immersion in aqua regia for about five minutes.

3. RESULTS

(a) *Analysis of the data*

Fig. 3 shows a plot of a set of measurements on the transition fields of the two specimens. Here the pressurized specimen H was under a pressure of 2080 psi, and consequently its critical field was about 1·5 gauss lower than that of the comparison specimen L. As explained before, we have plotted μ_e vs. H_e for each transition and extrapolated the straight line portion of the curve to $\mu_e = 1$ in order to obtain H_c. The critical fields are numbered in the order in which they were measured. The small successive displacements of the individual H and L transitions from one another are due to the temperature drift of the specimens during the measurements. The critical-field drift between transitions 1 and 5 corresponds to a temperature drift of 2×10^{-4}K°/hour.

VOL. 1 1956/5

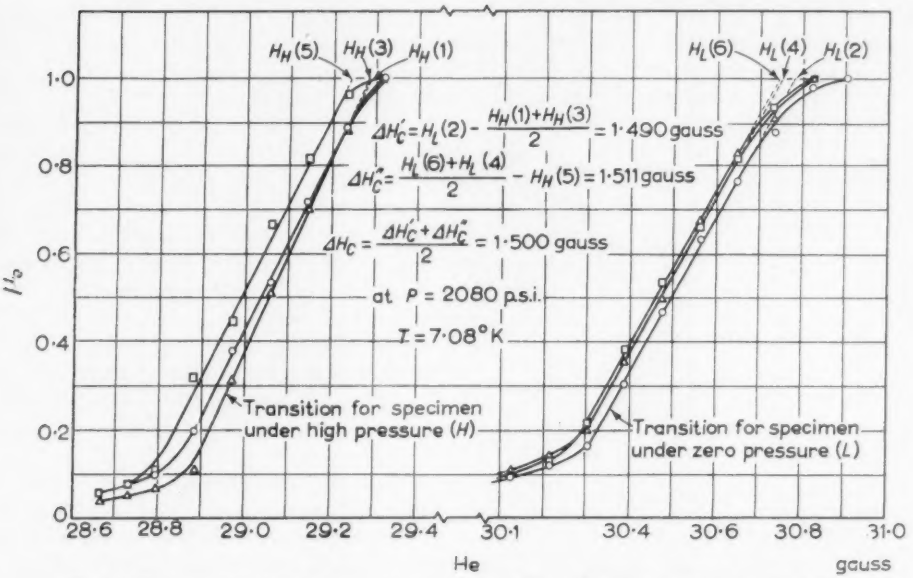

$H_H(5)$ $H_H(3)$ $H_H(1)$
$H_L(6)$ $H_L(4)$ $H_L(2)$
$\Delta H_C' = H_L(2) - \frac{H_H(1)+H_H(3)}{2} = 1{\cdot}490$ gauss
$\Delta H_C'' = \frac{H_L(6)+H_L(4)}{2} - H_H(5) = 1{\cdot}511$ gauss
$\Delta H_C = \frac{\Delta H_C' + \Delta H_C''}{2} = 1{\cdot}500$ gauss
at $P = 2080$ p.s.i.
$T = 7{\cdot}08°$K
Transition for specimen under high pressure (H)
Transition for specimen under zero pressure (L)
I/I_0
1·0
0·8
0·6
0·4
0·2
0
28·6 28·8 29·0 29·2 29·4 30·1 30·4 30·6 30·8 31·0
He
gauss

FIG. 3. Superconducting-to-normal transitions taken to determine the pressure shift in the critical field for a pressure of 2080 psi. The critical-field pressure shift is computed as shown so as to compensate for the temperature drift.

In comparing the critical fields for the H and L specimens in Fig. 3, we have averaged the results of two three-measurement comparisons in order to get the pressure shift ΔH_c. The temperature drift during each of these three-measurement comparisons was about 0·005 gauss or $2{\cdot}5 \times 10^{-5}$K°. Fourteen separate series of measurements such as that shown in Fig. 3 were made at different pressures. In Table 1 we have tabulated all the ΔH_c vs. P measurements. The points are numbered in the order in which they were taken. Note that the data were taken at increasing and decreasing pressures over several cycles of pressure, so that any irreversible pressure effects would have become apparent. Shown in Table 1 are the values of the critical-field drifts during three-measurement comparisons $\Delta H(D)$ and the residuals $\delta H(R)$. The residuals shown are the differences between the measured values of ΔH_c for each point, and the value of ΔH_e as given by the least-squares curve shown in Fig. 4.

We believe that the points 1, 5, and 6 reflect experimental error. These points were taken relatively soon after pressure or temperature had been established, and we suspect that thermal

Table 1. *Data for the Pressure Effect in Lead*

Point No.	Pressure (psi)	ΔH_c (gauss) Pressure shift	No. of Transitions measured	$\Delta H(D)$ (gauss) Critical-field temperature drift	$\delta H(R)$ (gauss) Residual
1*	440	−0·592	3	0·011	+0·236*
2	445	−0·405	4	0·011	+0·046
				0·012	
3	1439	−0·966	4	0·091	−0·051
				0·125	
4	2920	−1·928	11	0·000	−0·070
				0·000	
5*	1468	−0·817	7	0·076	−0·219*
6*	525	−0·211	6	0·03	−0·201*
7	530	−0·435	7	0·008	+0·02
8	3115	−2·059	5	0·017	−0·068
9	1397	−0·948	4	0·13	−0·041
				0·16	
10	4298	−2·971	8	0·023	+0·061
11	4253	−2·905	5	0·003	+0·025
12	2080	−1·500	6	0·005	+0·059
				0·005	
13	545	−0·474	7	0·005	+0·049
14	25	−0·051	6	0·010	−0·030

* Anomalous points.

equilibrium had not been attained in the inner can, resulting in a temperature difference on the order of 10^{-3}°K between the specimens. We have ignored these anomalous points in plotting Fig. 4 and in stating the final results.

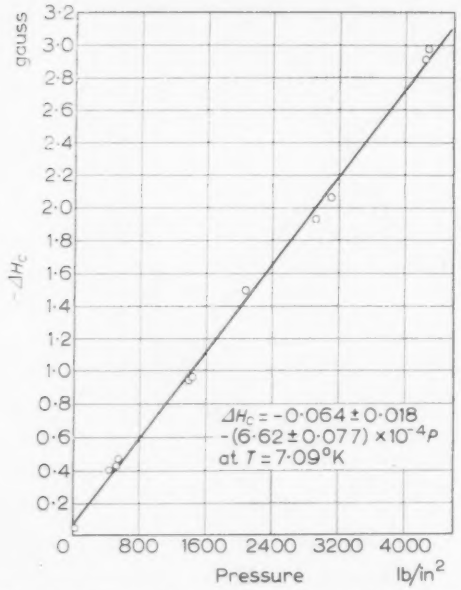


FIG. 4. Pressure shift in the critical field of lead at constant temperature.

In Fig. 4 we have plotted the ΔH_c vs. P data for the eleven good points. The least-squares curve is

$$\Delta H_c(T_m) = -(0{\cdot}064 \pm 0{\cdot}018) - (6{\cdot}62 \pm 0{\cdot}077) \times 10^{-4} P \quad (1)$$

which gives

$$(\partial H_c/\partial P)_{T_m} = -6{\cdot}62 \pm 0{\cdot}077 \times 10^{-4} \text{ gauss/psi} = -9{\cdot}60 \pm 0{\cdot}11 \times 10^{-9} \text{ gauss/dyne cm}^{-2} \quad (2)$$

where T_m is the average temperature of measurement, 7·09°K.*

The probable error in (2), as calculated from the residuals, is 1·2 per cent. The linearity of the ΔH_c vs. P curve is in agreement with previous results on In[7] and Sn.[7, 12, 13, 19]

The zero-pressure intercept of −0·064 gauss may be due to a slight temperature difference (about 3×10^{-4}K°) between the specimens throughout the course of the measurements or an inherent difference in the critical fields of the two specimens. There is no evidence that the displacement was due to an irreversible pressure effect.[13] The high-pressure specimen was cooled down to helium temperature under a pressure of 500 psi and point number 2 was taken at about this pressure. Subsequent cycling between high and low values (see Table 1) did not change the value of ΔH_c measured at the low pressures.

All the transitions measured to determine the pressure shift were taken in the superconducting to normal direction. Some other transitions taken in the normal to superconducting direction showed a slight degree of supercooling. The lowest supercool field observed at 7·09°K was about $0{\cdot}98 H_c$.

(b) *Critical-temperature shift*

Discussion of the pressure effect is facilitated by expressing the pressure coefficient in terms of the pressure derivative of the zero-field critical temperature, dT_c/dP. This quantity is best derived from critical-field shift measurements taken at constant temperature near the critical temperature T_c, and use of the relation†

$$dT_c/dP = -(\partial H_c/\partial P)_{T_c} / (\partial H_c/\partial T)_{T_c}.$$

Here $(\partial H_c/\partial P)_{T_c}$ can be obtained by measuring $(\partial H_c/\partial P)_T$ as a function of T and extrapolating the measured values to T_c. $(\partial H_c/\partial T)_{T_c}$ can be obtained by direct measurement of H_c vs. T near T_c at zero pressure. Due to experimental difficulties, however, neither $(\partial H_c/\partial P)_T$ nor H_c was measured as a function of temperature. A good estimate of dT_c/dP is possible, nevertheless. Since the temperature of measurement T_m was only about 0·13°K below T_c, $(\partial H_c/\partial P)_{T_m}$ and $(\partial H_c/\partial P)T_c$ would be expected to differ by a few per cent at most.

* This temperature is computed from the average zero pressure critical field, 29·75 gauss and the H_c vs. T relation given in eq. (3) of the next section. The maximum deviation from this temperature for any point shown in Fig. 4, as judged from the critical field of the zero-pressure comparison specimen, was about 0·002°K. Assuming the "similarity principle" this temperature deviation would be expected to cause changes in ΔH_c of only about 0·2 per cent.

† The subscript T_c is used here to signify that the partial derivatives are to be evaluated at the critical temperature. This usage differs from the usual thermodynamic notation but it is believed that the meaning is clear in this case.

VOL. 1 1956/57

$(\partial H_c/\partial T)_{T_c}$ can be obtained from RUTGER'S equation and the specific-heat data on lead of CLEMENT and QUINNELL.[14] Also H_c vs. T measurements have been made from 1·5°K up to 6·85°K by DAUNT, HORSEMAN, and MENDELSSOHN.[15, 16] Although there is scatter in the data above 4·2°K, a parabola

$$H_c = H_0[1-(T/T_c)^2] \tag{3}$$

with $H_0 = 805$ gauss[17] and $T_c = 7{\cdot}22$°K[18] appears to be a fair fit to the data above 5·8°K.

In order to estimate $(\partial H_c/\partial P)_{T_c}$ from $(\partial H_c/\partial P)_{T_m}$ we shall use (3) and assume that the "similarity principle" applies, i.e. $H_c/H_0 = f(T/T_c)$ independent of P, and $\partial(H_0/T_c)/\partial P = 0$.* This gives

$$(\partial H_c/\partial P)_{T_c} = \frac{2}{1+(T_m/T_c)^2}(\partial H_c/\partial P)_{T_m}. \tag{4}$$

Substituting $T_m = 7{\cdot}09$°K† and $T_c = 7{\cdot}22$°K[18] into (4) yields

$$(\partial H_c/\partial P)_{T_c} = 1{\cdot}019(\partial H_c/\partial P)_{T_m}$$

or a 2 per cent difference. The temperature dependence of the pressure effect has been measured for In and Sn. For In[7] and also for Sn, according to FISKE,[19] the decrease of $(\partial H_c/\partial P)_T$ with decreasing temperature is about half what would be calculated from the similarity principle. Because of this and because (3) may not be exactly correct, we guess that the error in the temperature-correction factor may be as high as 2 per cent. This possible error must be added to the error in $(\partial H_c/\partial P)_{T_m}$ for a total error in $(\partial H_c/\partial P)_{T_c}$ of 3·2 per cent, hence

$$(\partial H_c/\partial P)_{T_c} = -(6{\cdot}75 \pm 0{\cdot}20)\times 10^{-4}\,\text{gauss/psi}. \tag{5}$$

RUTGER'S equation is

$$(\partial H_c/\partial T)_{T_c} = (4\pi\Delta C/VT_c)^{\frac{1}{2}} \tag{6}$$

where $\Delta C = C_s - C_n$ is the molar-specific-heat difference between the superconducting and normal states at T_c, and V is the molar volume at T_c. For lead we have $V = 17{\cdot}8 \pm 0{\cdot}2$ cm³/mole, where the uncertainty is due to possible inaccuracy in correcting for the thermal contraction between room temperature and T_c. CLEMENT and QUINNEL[15] found $\Delta C = 12{\cdot}6 \pm 0{\cdot}5 \times 10^{-3}$ cal/mole °K. Substituting these values in (6) gives

$$(\partial H_c/\partial T)_{T_c} = 227 \pm 5{\cdot}7\,\text{gauss/°K} \tag{7}$$

in agreement with the value obtained by differentiation of (3) which is

$$(\partial H_c/\partial T)_{T_c} = 223\,\text{gauss/°K}.$$

The possible error in $(\partial H_c/\partial T)_{T_c}$, 2·5 per cent, must be added to the possible error in $(\partial H_c/\partial P)_{T_c}$ for a total possible error in dT_c/dP of 5·7 per cent. This gives

$$dT_c/dP = -(4{\cdot}31 \pm 0{\cdot}25)\times 10^{-11}\text{K°/dyne cm}^{-2}. \tag{8}$$

In discussing the results it will be convenient to express (8) in terms of the logarithmic volume deviative. Using an extrapolated value for the compressibility χ_0 of lead at T_c, $2{\cdot}06\times 10^{-12}$ dyne^{-1} cm², we obtain

$$d\ln T_c/d\ln V = 2{\cdot}89 \pm 0{\cdot}16. \tag{9}$$

4. DISCUSSION

An expression for the volume derivative of the critical temperature may be derived from the Bardeen theory. Although the theory is not exact enough to be quantitatively correct, it does suggest the factors which may be of importance in the pressure effect. According to BARDEEN,[20] the energy difference per unit volume between normal and superconducting states is

$$W_s - W_n = -H_0^2/8\pi = KN(E_F)a^2 \tag{10}$$

where K is a numerical constant independent of volume, $N(E_F)$ is the number of electronic energy states per unit energy range per unit volume at the Fermi level, and where a is related to the high-temperature relaxation time τ of conductivity theory and the Debye temperature θ_D, by the equation

$$a = h\theta_D/4\pi T\tau. \tag{11}$$

From conductivity theory[21] the usual expression for $T\tau$ is such that

$$T\tau \propto \theta_D^2 M/m^* C^2 \tag{12}$$

* These two conditions imply that γ, the coefficient of the electronic specific heat per unit volume in the normal state is independent of pressure.

† See footnote on page 205.

where $M \equiv$ atomic mass
$m^* \equiv$ effective electronic mass
$C \equiv$ the electron-phonon interaction constant which measures the strength of the scattering matrix elements.

Substituting (12) into (11) we have

$$a \propto m^* C^2 / M\theta_D. \tag{13}$$

Equation (10) can be written in terms of T_c, the critical temperature. Since all critical field curves are approximately parabolic, we write the reduced critical field, $h_c = H_c/H_0$, as a function of the reduced critical temperature $t = T/T_c$ of the general form

$$h_c = 1 - At^2 + Bt^3 + \ldots . \tag{14}$$

Then we have, from thermodynamics,[22]

$$\gamma = -\lim \frac{1}{8\pi} \frac{H_0^2}{T_c} \frac{d^2 h_c^2}{dt^2} = \frac{A}{2\pi} \left(\frac{H_0}{T_c} \right)^2 \tag{15}$$

where γ is the coefficient of the electronic specific heat per unit volume in the normal state. Substituting H_0 from (15) into (10) gives

$$T_c \propto N(E_F)^{\frac{1}{4}} a A^{\frac{1}{4}} / \gamma^{\frac{1}{4}}$$

or since $\gamma \propto N(E_F)$ this may be written, using (13)†

$$T_c \propto a A^{\frac{1}{4}} \propto m^* C^2 A^{\frac{1}{4}} / M\theta_D. \tag{16}$$

Differentiating (16) we obtain‡

$$\begin{aligned} d\ln T_c / d\ln V &= \eta_0 + d\ln m^*/d\ln V + 2 d\ln C / d\ln V \\ &\quad + \tfrac{1}{2} d\ln A / d\ln V \end{aligned} \tag{17}$$

where $\eta_0 = -d\ln \theta_D / d\ln V$ is GRÜNEISEN'S constant at T_c.[23]

If the last three factors on the right of (17) are neglected, then the expression reduces to a relationship first pointed out by MARCUS:[24]

$$d\ln T_c / d\ln V = \eta_0. \tag{18}$$

Here only the effect of zero-point lattice-vibration-amplitude changes are considered. For Sn, In, and Al (18) gives values of $d\ln T_c/d\ln V$ smaller than the measured values by factors of about 3·3, 2·5, and 6·0, respectively; and gives the wrong sign for the pressure coefficients of Bi_4Rh, Bi_3Ni, Bi_2K, and Tl (below 1500 atmospheres). It seems doubtful that such large discrepancies could result from the use of incorrect values of χ_0§ and η_0. On the basis of the Bardeen theory, the discrepancies between measured values of $d\ln T_c/d\ln V$ and (18) can be explained as resulting from the contribution of the last three "electronic" factors in (17). The situation here is analogous to the case of the volume dependence of the high-temperature resistivity. From conductivity theory an expression can be derived for $d\ln \rho/d\ln V$ which is similar to (17) and, like (17), contains factors which express the influence of vibration-amplitude changes and factors which express the influence of electronic changes (i.e. changes in m^* and C).‖ Large deviations of $d\ln \rho/d\ln V$ from those expected only on the basis of reduced lattice vibration amplitude under compression (e.g. for Li, Ca, Sr), as estimated from GRÜNEISEN'S constant, are usually ascribed to the influence of the electronic factors.[25] For the superconductors Sn, In, Al, and Tl the measured values of $d\ln \rho/d\ln V$ (but not $d\ln T_c/d\ln V$) are in fairly good agreement with what might be expected on the basis of vibration-amplitude changes alone.[27] Hence for these metals the influence of electronic factors on $d\ln \rho/d\ln V$ appears to be considerably smaller than on $d\ln T_c/d\ln V$.

It will be noted that $d\ln A/d\ln V$ and $d\ln \gamma/d\ln V$ (and hence $d\ln m^*/d\ln V$ since $\gamma \propto m^* V^{-\frac{1}{3}}$)

† Note that (16) is in agreement with the isotope effect results since $(m^* C^2 A^{\frac{1}{4}})$ is essentially independent of isotopic mass and $(M\theta_D)^{-\frac{1}{4}} \propto M^{-\frac{1}{4}}$ at constant volume.

‡ A somewhat similar relation has been derived by S. MASE, *Buseiron Kenkyu* **53**, 10 (1952). He has evidently expressed τ in terms of the high temperature resistivity ρ through the relation $\tau = m^*/\rho(N/V)e^2$, where N/V is the number of electrons per unit volume and e is the electronic charge. Making this substitution and proceeding as before one obtains $d\ln T_c/d\ln V = -1 - \eta_0 + d\ln \rho/d\ln V - d\ln m^*/d\ln V + \frac{1}{2} d\ln A/d\ln V$. Mase does not include the last two factors in his paper.

§ C. A. SWENSON *Phys. Rev.* **100**, 1607 (1955) has shown that neither indium nor thallium exhibits unusual mechanical behavior up to pressures of 10,000 atmospheres at temperatures down to 4·2°K.

‖ Actually the interaction constant, C, is averaged over the scattering angle and so may be sensitive to the form of the vibrational spectrum. It is therefore perhaps an oversimplification to regard C as solely an "electronic" factor.

VOL. 1 1956/5

can, according to (14) and (15) be obtained from precise measurements of the effect of pressure on the critical-field curve over a large temperature range. Their effect would be to produce deviation from the temperature dependence of the pressure coefficient as predicted by the similarity principle (see equation 4). Electronic effects have been observed in such a manner for Sn[19] and In.[7]

The present measurement suggests that in lead, unlike in other superconductors previously measured, the pressure-induced shift of the critical temperature is due primarily to the effect of pressure on the lattice vibrations, and that the net effect of electronic factors is small. From (18) we obtain

$$d\ln T_c/d\ln V = 2{\cdot}73 \qquad (19)$$

in good agreement with (9). Here it is assumed that GRÜNEISEN'S constant η is temperature independent so that η_0, the value at T_c, can be computed from high-temperature measurements of thermal expansion coefficient α, compressibility χ, molar heat capacity C_v, and the molar volume at absolute zero V_0 by substituting into the expression $\eta_0 = \eta = \alpha V_0/\chi C_v$.* Recent measurements on the thermal expansion coefficient of Al[26] and Cu[26, 27] at temperatures down to 20°K have shown that while η is constant above $0{\cdot}4\theta_D$, there is a decrease at temperatures below this to a value possibly 25 per cent less[26] at about $0{\cdot}1\theta_D$. This decrease has been explained by BARRON[28] in terms of BORN-VON KARMAN'S lattice dynamics by assuming central forces between atoms in an f.c.c. lattice. It is to be expected that at still lower temperatures η should become constant again but the measurements have not been extended to low enough temperatures to show this. If the decrease in η is real for Al and occurs also in Sn and In then the discrepancies between (18) and the measured values for $d\ln T_c/d\ln V$ are increased for these superconductors. Since Pb, like Al and Cu, is f.c.c., it is possible that its Grüneisen constant behaves similarly and also decreases by about 25 per cent at $T_c \approx 0{\cdot}1\theta_D$. If such were the case then (18) would yield a value for $d\ln T_c/d\ln V$ about 40 per cent smaller than the measured value so that (18) would still be in better agreement with experiment for Pb than for other superconductors.

It should be pointed out that, according to present ideas on lattice dynamics, volume or isotopic mass changes should produce changes in the shape of the vibrational spectrum of lead at temperatures in the vicinity of T_c.[29] This is related to the fact that lead, unlike most superconductors, has a value of T_c/θ_D of such magnitude (0·08) that departure from a Debye spectrum is to be expected. It is possible that the transition temperature is sensitive to the shape of the vibrational spectrum and DE LAUNAY[29] has argued that the unusually large shift reported for the isotope effect in lead[30, 31] ($T_c \propto M^{-0.7}$ instead of $T_c \propto M^{-0.5}$) is evidence for this.

Further measurements of the pressure shift and also the isotope shift in lead over a wide temperature range are now in progress in this laboratory.

Acknowledgements—We wish to thank Mr. R. S. KAESER for his preparation of the lead specimens and for his very able assistance in making the measurements. The advice of Mr. F. E. WITT on design and construction problems was most valuable. We are also grateful to Prof. J. BARDEEN and Dr. F. S. HAM for helpful discussions.

REFERENCES

1. MAXWELL E. *Phys. Rev.* **78,** 477 (1950).
2. REYNOLDS C. A., SERIN B., WRIGHT W. H., and NESBITT L. B. *Phys. Rev.* **78,** 487 (1950).
3. (*a*) MAXWELL E. *Phys. Today* **5,** No. 12, 14 (1952); (*b*) SERIN B. *Progress in Low Temperature Physics*, I, ed. by C. J. GORTER (Interscience Publishers, Inc., New York, 1955), p. 142.
4. BARDEEN J. *Phys. Rev.* **80,** 567 (1950).
5. FROHLICH H. *Phys. Rev.* **79,** 845 (1950).
6. SQUIRE C. F. *Progress in Low Temperature Physics*, I, ed. by GORTER C. J. (Interscience Publishers, Inc., New York, 1955), p. 151.
7. MUENCH N. L. *Phys. Rev.* **100,** 1784 (1955) gives a summary of published results.
8. ALEKSEEVSKI N. E. *J. Exp. Theor. Phys. U.S.S.R.* **19,** 358 (1949).
9. ALEKSEEVSKI N. E. and BRANDT N. B. *J. Exp. Theor. Phys. U.S.S.R.* **22,** 200 (1952).
10. HATTON J. *Phys. Rev.* **100,** 1784 (1955).
11. COCHRAN J. F., MAPOTHER D. E., and MOULD R. E. *Phys. Rev.* **103,** 1657 (1956).
12. GARBER M. and MAPOTHER D. E. *Phys. Rev.* **94,** 1065 (1954).
13. CHESTER P. F. and JONES G. O. *Phil. Mag.* **44,** 1281 (1953).
14. CLEMENT J. R. and QUINNELL E. H. *Phys. Rev.* **85,** 502 (1952).

15. DAUNT J. G. and MENDELSSOHN K. *Proc. Roy. Soc.* **A160,** 127 (1937).
16. DAUNT J. G., HORSEMAN A., and MENDELSSOHN K. *Phil. Mag.* **27,** 754 (1939).
17. SHOENBERG D. *Superconductivity* (Cambridge, 1952), p. 226.
18. BOORSE H. A., COOK D. B., and ZEMANSKY M. W. *Phys. Rev.* **78,** 635 (1950).
19. FISKE M. D. *Conference de Physique des Basses Temperature* (Centre National de la Recherche Scientifique, Paris, 1955), p. 518.
20. BARDEEN J. *Phys. Rev.* **80,** 567 (1950) eq. (4.26) p. 572.
21. MOTT N. F. and JONES H. *Properties of Metals and Alloys* (Oxford University Press, 1936).
22. MAXWELL E. *Phys. Rev.* **86,** 235 (1952).
23. SLATER J. C. *Introduction to Chemical Physics* (McGraw-Hill Book Company, Inc., New York, 1939) p. 219, 238, 450.
24. MARCUS P. M. *Phys. Rev.* **91,** 216 (1953).
25. LAWSON A. W. "The Effect of Hydrostatic Pressure on the Electrical Resistivity of Metals", to be published in *Progress in Metal Physics*, ed. by BRUCE CHALMERS (Interscience Publishers, Inc., New York, 1955).
26. BIJL D. and PULLAN H. *Physica* **21,** 285 (1955).
27. RUBIN T., ALTMAN H. W., and JOHNSTON H. L. *J. Am. Chem. Soc.* **76,** 5289 (1954).
28. BARRON T. H. J. *Phil. Mag.* **46,** 720 (1955).
29. DE LAUNAY J. *Phys. Rev.* **93,** 661 (1954).
30. OLSEN M. *Nature* **168,** 245 (1951).
31. SERIN B., REYNOLDS C. A., and LOHMAN C. *Phys. Rev.* **86,** 162 (1952).

VOL.
1
1956/5

J. Phys. Chem. Solids. Pergamon Press 1957. Vol. 1. pp. 209–227.

THE THEORY OF FERROMAGNETIC RESONANCE AT HIGH SIGNAL POWERS

H. SUHL

Bell Telephone Laboratories, Incorporated, Murray Hill, N.J.

(*Received* 4 *September* 1956)

Abstract—BLOEMBERGEN and WANG, as well as DAMON, have observed two anomalous effects in the microwave absorption of ferromagnets at high signal powers: a saturation of the usual resonance at powers far below those thought to be necessary on the basis of the observed relaxation time, and the appearance, at similar signal levels, of a rather broad secondary absorption peak some hundreds of oersteds below the d.c. field required for resonance. It is shown that these effects are connected with two kinds of instability of the uniform precessional motion of the total magnetization against certain spin-wave disturbances. These disturbances will grow, exponentially to begin with, when the signal level exceeds certain threshold values. It turns out to be possible to calculate the new final state attained by the medium in the case of the subsidiary absorption; in this state the uniform precession limits at a fixed value, no matter how large the signal, whereas a narrow range of spin waves is excited to a large nonthermal value which increases steadily with signal power. The susceptibilities calculated in the new regime check the experiments quite well. Some predictions of the theory beyond the above-mentioned observations have also been verified.

A simplified, less rigorous treatment has already been given elsewhere. In the present paper a more complete theory is developed, with special emphasis on the new state of the medium beyond the threshold.

1. INTRODUCTION

THE usual theory of ferromagnetic resonance is based on the concept that, as the result of the strong exchange forces between the spins, the spin system will precess with a high degree of spatial correlation in an applied microwave field. This is particularly evident in the case of an ellipsoidally shaped, single-crystal sample of a ferromagnetic insulator, placed into perfectly uniform fields. Leaving aside thermal agitation of the spin system, the motion of the magnetization in that case is perfectly uniform. The theory of this motion, already sketched by LANDAU and LIFSHITZ[1] in 1935, was perfected, in its phenomenological aspects, by KITTEL.[2] KITTEL'S theory confirmed on a microscopic basis by VAN VLECK,[3] predicted the position of the experimentally observed resonance line with considerable success. There remained only the question of the origin of the considerable width of the resonance line, and although some progress has recently been made towards a solution of this problem,[4, 5, 6] the precise mechanism is still not entirely clear in all cases.

KITTEL'S theory was based on a small signal solution of the equation of motion for the magnetization vector $\boldsymbol{M}$.

$$d\boldsymbol{M}/dt = -\gamma \boldsymbol{M} \times \boldsymbol{H}_{\text{eff}} + \text{a dissipative term} \qquad (1)$$

where γ is the absolute value of the gyromagnetic ratio of the electron-spins, and $\boldsymbol{H}_{\text{eff}}$ the total effective field, comprising, in the case of uniform motion, the applied steady and microwave fields, the shape demagnetizing field, and the crystalline anisotropy field. On the assumption of uniform motion, and with certain particular forms for the dissipative term, it is also possible to find the large signal solution of equation 1. In general this solution is very complicated, except in the case of spherical sample geometry and low crystalline anistropy. For such a sample one finds that the imaginary part, χ'', of the susceptibility at resonance essentially retains its small-signal value, until

the magnetic field of the signal becomes comparable with the line width of the resonance. Thereafter the χ'' declines steadily, and the decline is referred to as saturation of the resonance.

The first major conflict between observation and theory arose in the early experiments of DAMON[7] and in the more extensive measurements of BLOEMBERGEN and WANG.[8]* The original objective of these experiments has been to saturate the resonance in order to shed light on the relaxation processes, that is on the dissipative term in equation 1. However, two unexpected effects came to light:

1. The saturation of the usual resonance began at a signal power only about one hundredth of that required for saturation according to the conventional theory. In that power range, the simple linearized Kittel theory should still apply without appreciable change.
2. At a similar signal power, a secondary, rather broad absorption peak appeared a few hundred oersteds below the d.c. magnetic field required for the main resonance.

It will be shown in this paper that these effects arise from a spontaneous transfer of motion from the uniform precession to certain spatially fluctuating disturbances (spin waves). It turns out that the spin-wave motions are coupled in second and higher orders to the uniform precession by the demagnetizing fields and, to a less significant extent, by the exchange fields accompanying the spin waves. The coupling is such that certain of the spin waves abstract power from the uniform precession. At ordinary signal levels, this effect merely imposes a small additional loss on the uniform motion. Above a certain threshold signal, however, the power transfer becomes catastrophic, causing the relevant spin waves to grow with time, initially in an exponential manner. Their growth takes place at the expense of the uniform precession and causes the precession angle to diminish, until it attains a final value just below that appropriate to the threshold signal. By that time the spin waves have attained a large nonthermal value. This situation bears a certain resemblance to the turbulent state in fluid dynamics; fortunately however, the present problem is just sufficiently simple, so that the final state can actually be determined fairly rigorously, at least in case 2 of the B.W.-D. observations. In case 1, the algebraical complexity is still very great; nonetheless little doubt remains about the qualitative conclusions. For a simplified treatment of the problems here discussed, the reader is referred to ref. 9.

The idea that spin waves have some effect on the resonance phenomenon is not new. It has long been realized that the uniform precession is an independent, natural mode of motion only so long as the spin Hamiltonian is restricted to Zeemann and exchange terms. As soon as the dipolar interactions of the spins are added to the energy, the uniform precession is a natural mode only to first order; in second and higher orders more complicated modes are coupled to the uniform precession. In a certain approximation these higher modes have the character of plane waves and are coupled to the equation of motion of the uniform precession essentially as dissipative terms whose magnitudes depend on the amplitudes of the waves, and hence on the temperature, so long as thermal agitation is their sole driving mechanism. Line-width calculations based on such temperature dependent dissipative terms have been made by KASUYA[10] KEFFER,[11] and (without use of the spin-wave representation) by VAN VLECK.[3] (KEFFER's calculation does not actually yield a linewidth except under special conditions, but rather an off-resonance loss (see sections 5 and 6).)

What has been overlooked is the fact that just as the spin waves are coupled into the equation of motion of the uniform precession, the uniform precession is coupled into the equation of motion of the spin waves, and that this coupling can lead to the spectacular effects mentioned above.

For a number of reasons, the theory to be developed below will be semiclassical. Were one to proceed on a quantum-mechanical basis, one would obtain the equation of motion of the j^{th} spin, $\boldsymbol{S}_j$, of the assembly, by taking its commutator with the spin-Hamiltonian

$$\mathcal{H} = g\beta\Sigma\boldsymbol{H}\cdot\boldsymbol{S}_j + \Sigma A_{ik}\boldsymbol{S}_j\cdot\boldsymbol{S}_k + \\ + \sum\frac{g^2\beta^2}{r_{ik}^3}\left[(\boldsymbol{S}_i\cdot\boldsymbol{S}_k) - \frac{3}{r_{ik}^2}(\boldsymbol{S}_i\boldsymbol{r}_{ik})(\boldsymbol{S}_j\boldsymbol{r}_{jk})\right]$$

where g is the g-factor of the electrons, β is the Bohr magneton, and $\boldsymbol{H}$ the total applied field, steady and microwave. A_{ik} is the exchange constant between the

* Hereafter the papers in references 7 and 8 will be denoted by B.W.-D.

VOL. 1 1956/57

i^{th} and k^{th} ions, nonzero only if these are nearest neighbors, r_{ik} is the distance between ions i and k. Pseudomultipole terms have been neglected. They are concerned with anisotropy effects that do not seem to affect the B.W.-D. data in any important way. The resulting equations of motion can be written

$$\dot{\boldsymbol{S}}_j = -\gamma[\boldsymbol{S}_j \times \boldsymbol{H}_{\text{eff}}(j)] \; (j = 1, 2 \ldots N) \qquad (2)$$

in which $\boldsymbol{H}_{\text{eff}}$ (j) is the total effective field seen by spin j, which is the sum of the applied field, and of the exchange and dipolar fields that involve, respectively, the nearest neighbors of S_j, and all the remaining spins of the assembly. Suppose now that the d.c. field applied along Oz is large enough and the signal field and thermal agitation are small enough so that S_{jz} is almost constant and may be replaced by its maximum value S, and so that products of the transverse components of the various $\boldsymbol{S}_j$ are negligible. The right-hand side of (2), or rather the component-equation for the transverse parts of $\boldsymbol{S}_j$, can then be linearized, and yields a linear combination of the transverse components of all the $\boldsymbol{S}_i$. The "natural modes" of the system (2) may then be found; to a good approximation they turn out to be combinations of pairs of "spin waves", $\lambda_k S_k{}^+ + \mu_k S^-{}_{-k}$ where λ_k and μ_k are certain constants, and

$$S_k{}^+ = \sqrt{\frac{S}{N}} \sum_{j=1}^{N} S_j{}^+ e^{i\boldsymbol{k}\cdot\boldsymbol{r}_j};$$

$$S_{-k} = \sqrt{\frac{S}{N}} \sum_{j=1}^{N} S_j{}^- e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_j}$$

$$S_j{}^+ = S_{xj} + iS_{yj}$$

$$S_j{}^- = S_{xj} - iS_{yj}.$$

Here the $\boldsymbol{r}_j$ are the lattice sites, and $\boldsymbol{k}$ is any one of the N vectors of the reciprocal lattice times $1/N^{\frac{1}{3}}$. In particular, the uniform precession corresponds to $\boldsymbol{k} = 0$. Associated with each $\boldsymbol{k}$ value is a natural frequency ω_k. In the linear approximation the quantities $\lambda_k S_k{}^+ + \mu_k S^-{}_{-k}$ behave in all respects like harmonic oscillators with frequency ω_k. In the next and higher approximations, the right-hand side of (2) contains products of two transverse components, which couple the various spin waves together. Care must now be taken to place the correct interpretation upon

$$S_{zj} = \sqrt{S^2 + S - (S_{xj}{}^2 + S_{xj}{}^2)}$$

which can no longer be replaced by S, but must somehow be expanded in powers of $S_{xi}{}^2 + S_{xj}{}^2 - S$. However, the correct expansion procedure is still in doubt, both in the present language, and in the alternative formalism of HOLSTEIN and PRIMAKOFF in which the burden of the expansion is placed on the transverse spin components, rather than on S_z. (In particular, DYSON[15] has recently attacked the latter procedure as giving too great an interaction of long-wavelength spin waves when the Hamiltonian is restricted to Zeemann and exchange terms.) Disregarding these difficulties one would find, proceeding from equation 2, that the nonlinear coupling between certain of the spin waves is such that they fail to be stable normal modes. This will become clear later in the discussion. What is required to establish the spin waves as stable normal modes at least for small motions is a dissipative mechanism which tends to stabilize them. All that is known at present is that such mechanisms exist, but their precise nature is not fully understood. Therefore, it is necessary to supplement equation (2) by an entirely hypothetical damping term. It then becomes questionable if the quantum-mechanical approach to the problem is of any more value than a semiphenomenological treatment, in which one speaks in terms of the magnetization per unit volume, $\boldsymbol{M}(\boldsymbol{r})$ at the point $\boldsymbol{r}$ rather than in terms of the spin S_j at lattice position r_j.

Equation (2) is then replaced by

$$d\boldsymbol{M}(\boldsymbol{r}, t)/dt = -\gamma[\boldsymbol{M}(\boldsymbol{r}, t) \times \boldsymbol{H}_{\text{eff}}(\boldsymbol{r}, t)] + \\ + \text{a dissipative term.} \qquad (3)$$

$\boldsymbol{H}_{\text{eff}}$, as before, consists of the applied field, the exchange field, and the dipolar (or "demagnetizing") field, the last now calculable on the basis of classical continuum magnetostatics (electromagnetic propagation effects are negligible here). The semiclassical approach is justified only at temperatures sufficiently below the Curie point, yet not so low that zero-point motion is important. For, disregarding the dissipative term, equation (3) demands that $|\boldsymbol{M}|$ be a constant of the motion. Since $\boldsymbol{M}$ is defined as an average over many individual spin magnetic moments, that requirement is fulfilled only if the spin direction does not vary too rapidly from site to site, that is to say, the important spin waves must be many lattice spacings long. This will be true at low temperatures as far as thermal excitation is concerned, and it is likewise true of the spin waves that grow at the expense of the uniform precession.

For definiteness in some of the computations, the dissipative term in equation (3) has been taken in the Landau-Lifshitz form

$$\frac{\alpha}{|M|}[\boldsymbol{M} \times \dot{\boldsymbol{M}}]$$

where α is a dimensionless constant. More generally, however, we shall assume the frequencies of the spin waves that are the first order solutions of

equation (3) to be slightly complex; the imaginary parts, the decay constants, depending on k in a still unknown manner. It is possible that comparison of future measurements with the theory presented here will eventually lead to a determination of this functional dependence, at least over a certain range of wavelengths.

2. EQUATIONS OF MOTION FOR THE SPIN-WAVE AMPLITUDES

The magnetization vector $\boldsymbol{M}(\boldsymbol{r}, t)$ at a point $\boldsymbol{r}$ in the sample may be expanded into a spatial Fourier series, whose coefficients are functions of the time t:

$$\boldsymbol{M}(\boldsymbol{r}, t) = \sum_{k} \boldsymbol{M}_k(t) e^{i\boldsymbol{k}\cdot\boldsymbol{r}}. \tag{4}$$

Since we are committed to the continuum approach, it would be more appropriate to use a Fourier integral representation; however, the series is easier to handle formally, and it is clear that no serious error is incurred if we allow the summation to extend over the reciprocal lattice vectors times $N^{-1/3}$ as in the discrete case. From (4) it follows immediately that

$$\boldsymbol{M}_k(t) = \frac{1}{V}\int_V \boldsymbol{M}(\boldsymbol{r}, t) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\, dv$$

the integration extending through the volume V of the sample. We must now express the total effective field, $\boldsymbol{H}_{\text{eff}}(\boldsymbol{r}, t)$ in terms of the applied field and in terms of the $\boldsymbol{M}_k$. Suppose that the sample is a spheroid of revolution about Oz, the direction of the applied steady field $\boldsymbol{H}_0$, and that the signal field $\boldsymbol{h}$ is circularly polarized in the x, y, plane, having components $h\cos\omega t$, $h\sin\omega t$ along x and y respectively.

The exchange contribution to $\boldsymbol{H}_{\text{eff}}$ has the form

$$\boldsymbol{H}_{\text{ex}} = H_e \frac{l^2}{|\boldsymbol{M}|}\nabla^2 \boldsymbol{M}(\boldsymbol{r}, t) \tag{5}$$

where H_e is a well-known equivalent field depending on the exchange constant, lattice spacing l and on S. The manner in which this field may be derived from the expression for the exchange energy has been described in detail by HERRING and KITTEL.[12] In terms of the Fourier expansion, we have

$$\boldsymbol{H}_{\text{ex}} = -H_e \frac{l^2}{M}\sum k^2 \boldsymbol{M}_k(t) e^{i\boldsymbol{k}\cdot\boldsymbol{r}}.$$

Evaluation of the demagnetizing field $\boldsymbol{H}^{\text{d}}$ is more difficult. As far as it is due to components of M with sufficiently large $|\boldsymbol{k}|$ (exceeding a minimum $k_{\min}$, equal to perhaps ten times the reciprocal sample dimensions), it is readily calculated simply by disregarding the boundary conditions at the air-sample interface. For, if $\boldsymbol{k}$ is large enough, the pole distribution induced by $\boldsymbol{M}_k e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ at the intersection of the planes of equal phase with the boundary alternates in sign very rapidly, and the fields arising from it therefore do not reach any appreciable distance into the interior of the sample. Neglecting, then, the boundary conditions, the demagnetizing field $\boldsymbol{H}^{\text{d}}{}_{k>k\min}$ immediately follows from the requirement that the divergence of the total flux in the sample must vanish, and from the fact that propagation effects are negligible:

$$\boldsymbol{H}^{\text{d}}_{k>k_{\min}} = -4\pi \sum_{k>k_{\min}} \frac{\boldsymbol{k}(\boldsymbol{k}\cdot\boldsymbol{M}_k)}{k^2} e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$$

$$\equiv \Sigma\, h_k^{(\text{d})} e^{i\boldsymbol{k}\cdot\boldsymbol{r}}.$$

The contribution to $\boldsymbol{H}^{\text{d}}$ from wave vectors smaller than $k_{\min}$ is much more difficult to calculate. The end effects are then important, as is especially evident in the classical case $\boldsymbol{k} = 0$ (uniform magnetization). Then we know from elementary magnetostatics that the demagnetizing field has components

$$\boldsymbol{h}_0{}^{\text{d}} = -4\pi(N_x M_{x0}, N_y M_{y0}, N_z M_{z0}) \tag{6}$$

where N_x, N_y, N_z are the demagnetizing factors, defined so that $N_x + N_y + N_z = 1$. (In the case of the spheroid of revolution about Oz, $N_x = N_y = N_T$.)

For intermediate values of $\boldsymbol{k}$, $\boldsymbol{h}_k{}^{(\text{d})}$ is a very complicated function of many $\boldsymbol{M}_{k'}$. This means that spin waves whose wavelengths are comparable with the sample dimensions are not normal modes of the system, even to first order. To find the correct modes it is necessary to solve equation (3), the demagnetizing field being determined from

$$\boldsymbol{H}^{\text{d}} = \nabla\phi; \quad \nabla^2\phi = -4\pi\nabla\cdot\boldsymbol{M} \tag{7}$$

and from the boundary conditions that the tangential

VOL. 1 1956/57

components of H^d and the normal component of $H^d+4\pi M$ be continuous across the sample-air interface. If the exchange term is included in (3) further boundary conditions are needed (the Rado-Herring conditions on $\nabla_n M$).[5] However, for the slow spatial rates of variation for which boundary conditions need be considered at all, the exchange term can be omitted. In that case the linearized form of equations (3) and (7) has been solved, subject to the magnetostatic boundary conditions, by WALKER.[13]

The problems considered here are not strictly linear, but the linearized solutions still form the starting point of the investigation. Some of the problems involve rapid spatial variation of M and the plane wave solutions are then adequate lowest-order solutions; others, however, involve long spin waves and then WALKER'S modes should form the starting point. It is, in fact, possible to write down the formal solution of these latter problems in terms of WALKER'S modes, but progress beyond the formal solution is not possible at present, since the results involve integrals as yet unknown over combinations of the wave functions replacing the $e^{i\mathbf{k}\cdot\mathbf{r}}$. Hence we shall use the spin-wave expression for the demagnetizing field even for $\mathbf{k}$ numbers for which it is not correct; we shall assume that (6) holds for $k=0$, and that

$$h_k^{(d)} = -4\pi\frac{\mathbf{k}(\mathbf{k}\cdot\mathbf{M}_k)}{k^2} \tag{8}$$

for $k>0$, no matter how small. Judging from the good agreement between theory and experiment no serious error seems to be incurred thereby, for reasons which will become clearer in section 3.

We are now ready to establish the equation of motion for the M_k from equations (3) to (6) and (8). Since equation (3) preserves M, one of the three component equations, say the one for M_z, is redundant. Where M_z occurs in the remaining two, we replace it by

$$M_z = \sqrt{(M^2-M_x^2-M_y^2)}. \tag{9}$$

Provided the temperature is low enough, so that the $\mathbf{M}_k$ are not too heavily excited, and provided the uniform precession (M_{x0}, M_{y0}) is not excessive, the square root may be expanded in powers of $M_x^2+M_y^2$, and M is to be interpreted as the saturation magnetization at absolute zero.

Although we thus restrict our treatment to low temperatures, the restriction is probably not very significant. Since only a certain group G of spin waves ever becomes important, one can in some sense write

$$M_x^2+M_y^2 =$$

$$(M_x^2+M_y^2)_{G+\text{uniform precession}}+(M_x^2+M_y^2)_{\text{therma}}$$

and then M^2 in (9) is replaced by the square of an effective saturation magnetization

$$M^2-(M_x^2+M_y^2)_{\text{thermal}} = M^2_{\text{thermal}}$$

at the finite temperature. It is, however, not easy to make this argument precise, since this decomposition of $M_x^2+M_y^2$ is possible only as a time-average.

It will be convenient to use the following notation

$$\gamma H_0 = \omega_H; \quad 4\pi\gamma M = \omega_M; \quad \gamma H_e = \omega_{ex}$$

$$\gamma h = \omega_s; \quad \mathbf{M}/M = \mathbf{m}; \quad \gamma \mathbf{h}^{(d)} = \boldsymbol{\omega}_d$$

$$m^+ = m_x+im_y$$

$$m^- = (m^+)^* = m_x-im_y$$

$$m_z = \sqrt{(1-m^+m^-)} \doteq 1-\tfrac{1}{2}m^+m^-$$

$$a_k = m_{kx}+im_{ky}$$

and, since m_x and m_y are real,

$$a_{-k}{}^* = m_{kx}-im_{ky}.$$

(m_{kx} and m_{ky} are not generally real.)

In terms of this notation, if i times the y-component equation (3) is added to the x-component equation, we find

$$\dot{m}^+ = \Sigma\dot{a}_k e^{i\mathbf{k}\mathbf{r}} = im^+[\omega_H+\omega_{dz}+\omega_{ex}l^2\nabla^2 m_z]-$$

$$-im_z[\omega_s e^{i\omega t}+\omega_{ex}l^2\nabla^2 m^+ +\omega_{dx}+i\omega_{dy}].$$

Noting that $m_z = 1-\frac{1}{2}\sum_{kk'} a_{k'}a^*_{k'-k}e^{i\mathbf{k}\cdot\mathbf{r}}$, rearranging the various product series on the right-hand side and equating coefficients of $e^{i\mathbf{k}\cdot\mathbf{r}}$, we get, for $k\neq 0$

$$-i\dot{a}_k = A_k a_k + B_k a_{-k}{}^* -$$

$$-\frac{\omega_M}{2}\sum_{k'\neq 0} a_{k-k'}[a_{k'}k_{T'}{}^* + a_{-k'}{}^* k_{T'}]\frac{k'_z}{k'^2} +$$

$$+\frac{\omega_M}{2}[N_z a_k \sum_{k'} a_{k'}a_{k'}{}^* - N_T a_0 \sum a_{k'}a_{k'-k}{}^*] +$$

$$+\frac{1}{2}\sum a_{k'}a_{k''}a_{k'+k''-k}^{*}\times$$

$$\left[\omega_{ex}l^2(k^2-2k\cdot k')+\omega_M\frac{(k_z-k'_z)^2}{|k-k'|^2}-\frac{\omega_M|k'_T|^2}{2}\frac{}{k'^2}\right]-$$

$$-\frac{\omega_M}{4}\sum_{k'\neq 0}a_{-k'}^{*}a_{k''}a_{k'}^{*}{}_{+k''-k}\frac{(k'_T)^2}{k'^2} \qquad (10)$$

where $k_T = k_x + ik_y$, and where all indeterminate contributions are to be omitted from the sums involving ω_M.

The constants A_k and B_k are given by

$$\left.\begin{aligned} Ak &= \omega_H-\omega_M N_z+\omega_{ex}l^2k^2+\omega_M\frac{|k_T|^2}{2k^2} \\ B_k &= \frac{\omega_M}{2}\frac{k_T{}^2}{k^2}. \end{aligned}\right\} \qquad (11)$$

On the other hand, for $k = 0$

$$-i\dot{a}_0 = a_0[\omega_H-\omega_M(N_z-N_T)]-\omega_s e^{i\omega t}-$$

$$-\frac{\omega_M}{2}\sum_{k\neq 0}a_{-k}(a_k k_T{}^{*}+a_{-k}{}^{*}k_T)\frac{k_z}{k^2}+$$

$$+\frac{\omega_M}{2}(N_z-N_T)a_0\sum a_k a_k{}^{*}+$$

$$+\frac{\omega_M}{2}\sum_{k\neq 0}a_k a_{k'}a_{k+k'}{}^{*}\left[\frac{k_z{}^2}{k^2}-\frac{1}{2}\frac{|k_T|^2}{k^2}\right]-$$

$$-\frac{\omega_M}{4}\sum_{k\neq 0}a_{-k}{}^{*}a_{k'}a_{k+k'}\frac{k_T{}^2}{k^2}. \qquad (12)$$

The fact that the equation for a_0 is not entirely a special case of that for a_k is, of course, due to the difference in the form of the demagnetizing fields $h_k{}^{(d)}$ and $h_0{}^{(d)}$, and to the presence of the driving field.

In the linear approximation, equation (10) is

$$-i\dot{a}_k = A_k a_k+B_k a_{-k}{}^{*} \qquad (13)$$

and represents two coupled harmonic oscillators a_k, $a_{-k}{}^{*}$. By a suitable transformation, we can find the two normal modes of this pair, and the corresponding frequencies. The non-linear terms in (10) can be grouped according to whether they do or do not contain one of the quantities a_k, $a_{-k}{}^{*}$. Those that do not, can, to a first approximation, be looked upon as driving terms that merely contribute to the thermal excitation of a_k, $a_{-k}{}^{*}$. Those nonlinear terms that do contain either of these quantities, since they also will contain other, time-dependent a's, effectively modulate A_k and B_k. But A_k and B_k determine the frequency of the normal modes, hence the nonlinear terms in question will cause a frequency modulation. As is well known from harmonic oscillator theory, such frequency modulation can drive the oscillator unstable under suitable circumstances, and this is also the case here. Let us now suppose that the sample is in its conventional state of motion, in which the only large oscillator amplitude is the uniform precession a_0, induced by the signal field, while all the other a_k's have only their small thermal values. Then we can further restrict the choice of the important nonlinear terms to those that, besides a_k or $a_{-k}{}^{*}$, contain the largest possible number of a_0's or $a_0{}^{*}$'s. As far as the bilinear terms in (10) are concerned, this selection takes place automatically: double terms involving a_k or $a_{-k}{}^{*}$ of necessity also involve a_0 or $a_0{}^{*}$. This is not the case for the triple terms in (10), of which we shall retain only those involving a_0 or $a_0{}^{*}$ twice. Equation (10) is thus replaced by

$$-i\dot{a}_k = A_k a_k+B_k a_{-k}{}^{*}-$$

$$-\omega_M\left(\frac{1}{2}(a_0 k_T{}^{*}+a_0{}^{*}k_T)\frac{k_z}{k^2}a_k+\frac{k_z k_T}{k^2}a_0 a_{-k}{}^{*}\right)+$$

$$+\frac{\omega_M}{2}\left(N_z-N_T+\frac{k_z{}^2-\frac{1}{2}|k_T|^2}{k^2}\right)|a_0|^2 a_k+$$

$$+\frac{1}{2}\left[\left(\omega_{ex}l^2k^2-N_T\omega_M+\omega_M\frac{k_z{}^2}{k^2}\right)\times a_0{}^2-\right.$$

$$\left.-\frac{\omega_M}{2}\frac{k_T{}^2}{k^2}|a_0|^2\right]a_{-k}{}^{*}. \qquad (14)$$

So long as we are only concerned with the free motion of the a_k, we can discard the remaining nonlinear terms altogether. It is interesting to note that (14) could have been derived more simply if we had assumed at the outset that a_0, a_k, $a_{-k}{}^{*}$ are the only excitations present.[9] Such a procedure is indeed sufficient for a derivation of the instability criteria, but it does not suffice for the calculation of the final state, which involves

VOL. 1 1956/5

all spin waves participating in the instability, not just one pair a_k, a_{-k}^*.

We now "solve" equation (14) as follows: We first diagonalize equation (13) by means of the following transformation (essentially the same as that used by HOLSTEIN and PRIMAKOFF[16] in diagonalizing their spin Hamiltonian):

$$\left.\begin{aligned} a_k &= \lambda_k b_k - \mu_k b_{-k}^* \\ a_{-k}^* &= -\mu_k^* b_k + \lambda_k b_{-k}^* \end{aligned}\right\} \tag{15}$$

where

$$\lambda_k = \cosh\frac{\psi_k}{2} = \lambda_{-k};$$

$$\mu_k = \sinh\frac{\psi_k}{2} e^{2i\phi_k} = \mu_{-k};$$

$$\tan\phi_k = \frac{k_y}{k_x}; \quad \tanh\psi_k = \frac{|B_k|}{A_k};$$

$$A_k = \omega_k \cosh\psi_k;$$

$$B_k = \omega_k e^{2i\phi_k}\sinh\psi_k$$

$$\begin{aligned} \omega_k^2 = A_k^2 - |B_k|^2; \\ = (\omega_H - N_z\omega_M + \omega_{ex} l^2 k^2)\times \\ \times(\omega_H - N_z\omega_M + \omega_{ex} l^2 k^2 + \omega_M \sin^2\theta) \end{aligned}$$

$$\sin\theta = \frac{|k_T|}{k} \tag{16}$$

When (15) is substituted into (13), it is found that

$$\dot{b}_k^* = i\omega_k b_k; \quad \dot{b}_{-k}^* = -i\omega_k b_{-k}^*.$$

The relation of ω_k to $\boldsymbol{k}$ is known as the dispersion relation, or the spin-wave spectrum. It will be further discussed in section 5. At this point we introduce the loss parameter η_k of the $\boldsymbol{k}^{\text{th}}$ spin wave by assuming ω_k to be slightly complex:

$$\omega_k \to \omega_k + i\eta_k = \Omega_k$$

so that

$$\dot{b}_k = i\Omega_k b_k; \quad \dot{b}_{-k}^* = -i\Omega_k^* b_{-k}^*.$$

Thus b_k and b_{-k}^* are damped harmonic oscillators and the general solution of (13) can be written

$$a_k = \lambda_k b_k^\circ e^{(i\omega_k - \eta_k)t} - \mu_k b_{-k}^{\circ *} e^{-(i\omega_k + \eta_k)t} \tag{17}$$

where b_k°, $b_{-k}^{\circ *}$ are complex constants. In the presence of the non-linear terms we still take the solution in the form (17), but now allow b_k°, $b_{-k}^{\circ *}$ to become functions of the time. This procedure eliminates the linear part of (14) altogether. Substituting (17) in (14), writing down the complex conjugate equation with $\boldsymbol{k}$ changed to $-\boldsymbol{k}$, and eliminating $b_{-k}^{\circ *}$ between the resulting two equations with the help of (16), we obtain the following equation for the variation of the amplitude:

$$\begin{aligned} \dot{b}_k^\circ = i(D_1 + T_1)\times \\ \times[(\lambda_k^2 + \mu_k\mu_k^*)b_k^\circ - 2\lambda_k\mu_k b_{-k}^{\circ *} e^{-2i\omega_k t}] - \\ -i[(D_2 + T_2)\lambda_k\mu_k^* + (D_2^* + T_2^*)\lambda_k\mu_k]b_k^\circ + \\ +i[(D_2 + T_2)\lambda_k^2 + (D_2^* + T_2^*)\mu_k^2]b_{-k}^{\circ *} e^{-2i\omega_k t} \end{aligned} \tag{18}$$

where the D's derive from the double terms of (14):

$$D_1 = -\frac{\omega_M}{k^2}[k_T^* a_0 + k_T a_0^*];$$

$$D_2 = -\omega_M \frac{k_z k_T}{k^2} a_0 \tag{19}$$

and the T's from the triple terms:

$$\left.\begin{aligned} T_1 &= \frac{|a_0|^2}{2}\left[\omega_M(N_z - N_T) + \omega_M\frac{k_z^2 - \frac{1}{2}|k_k|^2}{k^2}\right] \\ T_2 &= \frac{a_0^2}{2}\left[\omega_{ex} l^2 k^2 - \omega_M N_T + \omega_M\frac{k_z^2}{k^2}\right] - \\ &\quad -\frac{\omega_M}{4}|a_0|^2\frac{k_T^2}{k^2}. \end{aligned}\right\} \tag{20}$$

So long as we merely examine the stability of the usual steady state, the b_k will all be small compared with a_0, and if a_0 itself is not too large, it is permissible to substitute the linearized solution of (12) in equation (18). This linearized solution is

$$a_0 = \frac{\omega_s e^{i\omega t}}{(\omega_{res} - \omega) + i\eta_0} \equiv a_0^\circ e^{i\omega t} \tag{21}$$

with

$$\omega_{res} = \omega_H - \omega_M(N_z - N_T),$$

η_0 is the loss parameter of the uniform precession, equation (12) having been supplemented by the dissipative term $i\eta_0 a_0$.

Equation (18) and its "adjoint" (the complex conjugate equation with $-\boldsymbol{k}$ replacing $\boldsymbol{k}$) form a pair of linear equations with time-dependent coefficients. Their most general solution must of necessity be very complicated. Therefore we are led to examine various trial solutions. The simplest, and apparently the most important kind is one in

which b_k° and $\dot{b}_k^\circ$, are slowly varying functions of the time. For such a solution we may neglect all terms on the right-hand side of (18) that have rapidly varying coefficients, since they will only add a small, high-frequency ripple to the slowly varying b_k. The only terms in (18) that can vary slowly are those whose coefficients are constants, and those whose total time-variation is as $e^{i(\omega-2\omega_k)t}$ or as $e^{2i(\omega-\omega_k)t}$. The last two can vary slowly provided one of the equations $\omega = 2\omega_k$ or $\omega = \omega_k$ can be approximately satisfied for some k. When all other terms are neglected, we finally obtain

$$\dot{b}_k^\circ = i\omega_M\lambda_k(\mu_k k_T^* - \lambda_k k_T)\frac{k_z}{k^2}a_0^\circ b_{-k}^{\circ *}e^{i(\omega-2\omega_k)t} + \\ + i\Big[T_1(\lambda_k^2+|\mu_k|^2) + \\ + \frac{\omega_M}{4}\lambda_k\frac{\mu_k k_T^{*2}+\mu_k^* k_T^2}{k^2}|a_0^\circ|^2\Big]b_k^\circ + \\ + i\frac{\lambda_k^2}{2}\left(\omega_{ex}l^2k^2 - \omega_M N_T + \omega_M\frac{k_z^2}{k^2}\right)\times \\ \times a_0^{\circ 2}b_{-k}^{\circ *}e^{2i(\omega-\omega_k)t}. \qquad (22)$$

We shall now examine the solution of this equation in the two cases in which k is such that $\omega \sim 2\omega_k$, and such that $\omega \sim \omega_k$. The first of these involves a coupling term of order a_0°, the latter involves coupling of order $a_0^{\circ 2}$. That is to say, spin waves for which $\omega = 2\omega_k$ are inherently more tightly coupled to the uniform precession, and we therefore begin with these.

3. THE SUBSIDIARY ABSORPTION

Restricting equation (22) to terms of first order in a_0°, we obtain, for spin waves with $2\omega_k$ close to ω,

$$\dot{b}_k^\circ = \rho_k a_0^\circ b_{-k}^{\circ *}e^{i(\omega-2\omega_k)t}$$

where

$$\rho_k = i\omega_M\frac{\lambda_k(\mu_k k_T^* - \lambda_k k_T)k_z}{k^2} = \rho_{-k} \\ = i\omega_M\frac{\lambda_k(|\mu_k|-\lambda_k)k_z|k_T|e^{i\phi_k}}{k^2} \qquad (23)$$

from (16). Writing

$$b_k^\circ = u_k e^{i[(\omega/2)-\omega_k]t}$$

we obtain

$$\left[\frac{d}{dt}+i\left(\frac{\omega}{2}-\omega_k\right)\right]u_k = \rho_k a_0^\circ u_{-k}^*.$$

Eliminating u_{-k}^* from this and the "adjoint" equation

$$\left[\frac{d}{dt}-i\left(\frac{\omega}{2}-\omega_k\right)\right]u_{-k}^* = \rho_k^* a_0^{\circ *}u_k,$$

we find

$$\left[\frac{d^2}{dt^2}+\left(\frac{\omega}{2}-\omega_k\right)^2-\rho_k\rho_k^*|a_0^\circ|^2\right]u_k = 0. \qquad (24)$$

If $\rho_k\rho_k^*|a_0|^2 > [(\omega/2)-\omega_k]^2$, this equation will have an exponentially increasing solution, with growth constant $\sqrt{\rho_k\rho_k^*|a_0^\circ|^2-[(\omega/2)-\omega_k]^2}$, but if the actual spin wave a_k is to increase exponentially, this growth constant must exceed the decay constant η_k. We thus require for instability

$$|a_0^\circ|^2 > \left[\eta_k^2+\left(\frac{\omega}{2}-\omega_k\right)^2\right]/\rho_k\rho_k^* \qquad (25)$$

which establishes the threshold value of the precession amplitude $|a_0^\circ|$. It can be translated into a threshold for the signal, using (21). Instability is seen to occur when $\omega_s > \omega_{s\,\mathrm{crit}}$, with

$$\omega_{s\,\mathrm{crit}} = \\ = \frac{\sqrt{\left[\eta_k^2+\left(\frac{\omega}{2}-\omega_k\right)^2\right]\left[\eta_0^2+(\omega_{\mathrm{res}}-\omega)^2\right]}}{\rho_k\rho_k^*} \\ = \frac{2\omega_k\sqrt{\left[\eta_k^2+\left(\frac{\omega}{2}-\omega_k\right)^2\right]\left[\eta_0^2+(\omega_{\mathrm{res}}-\omega)^2\right]}}{\omega_M\sin\theta_k\cos\theta_k[\omega_k+\omega_H-\omega_M N_z+\omega_{ex}l^2k^2]} \qquad (26)$$

where we have used (23) and the relations (16).

$\omega_{s\,\mathrm{crit}}$ is of order ω^2/ω_M, that is to say far beyond normally attainable values of the signal field unless one of the following conditions holds.

(a) $\omega_k \sim \omega/2$; ω_{res} not close to ω
(b) $\omega_k \sim \omega/2$; $\omega_{\mathrm{res}} \sim \omega$
(c) ω_k not close to $\omega/2$; $\omega_{\mathrm{res}} \sim \omega$.

VOL. 1 1956/5

In case (a), the right-hand side of (26) is of order $\eta(\omega/\omega_M)$ (in actual fact between 0·3 and 0·8 of this value), which is more readily attainable. In case (b), the threshold is even lower, of order η^2/ω_M. Case (c) is ruled out, since it violates the condition on which the present solution is based, namely that only terms with slowly varying coefficients are to be retained in equation (18).

Case (a) corresponds to the commonly found experimental situation. Usually experimental conditions are such that the equation $2\omega_k = \omega$ can be satisfied by a real $\boldsymbol{k}$ number only for d.c. fields some hundreds of oersteds below the field required for resonance, so that ω_{res} is not close to resonance. To find the least possible threshold signal, and the corresponding d.c. field at which the effect will first be noticed as the signal is increased from zero, it is necessary to minimize (26) with respect to $\boldsymbol{k}$ and the d.c. field H_0, (or ω_H). Since the coefficient of the radicand in (26) is almost everywhere a slowly varying function of $\boldsymbol{k}$, whereas the radicand itself has a deep minimum at $2\omega_k = \omega$, the values of $\boldsymbol{k}$ and ω_H that minimise (26) must satisfy

$$2\omega_k = \omega. \tag{27}$$

If (26) is minimized subject to (27) one finds that ω_H is very close to the maximum in the subsidiary absorption reported in the B.W.-D papers (see Table 1 and Fig. 1). It is therefore reasonable to suppose (and it will be demonstrated in section 6) that the subsidiary absorption can be ascribed to an outbreak of instability among spin waves satisfying $\omega_k = \omega/2$. In the identification in Table 1 of the value of ω_H that minimizes (23) with the position of the maximum in the subsidiary absorption, we have anticipated that this maximum occurs close to the value of the d.c. field at which instability will occur for the least signal. The theory of the new steady state confirms this assumption. Further, in the actual computation, the Landau-Lifshitz form of the damping term has been assumed for definiteness. In that case it turns out that

$$\eta_0 = \alpha\omega; \quad \eta_k = \alpha A_k.$$

The results are of course not very sensitive to the choice of damping terms, especially as far as the location of the subsidiary peak is concerned.

Table 1. Comparison of predicted and measured ratio of field at subsidiary resonance to the field required for the main resonance

	Theory	Cavity Experiments of Bloembergen and Wang
Ni-Ferrite Sphere $4\pi M \approx 3320$ oe	0·83	0·73
Mn-Ferrite Disc $4\pi M \approx 3880$ oe		
H∥ Disc	0·78	0·66
H⊥ Disc	0·72	0·72

$4\pi M$		Theory	Waveguide Experiments of M. T. Weiss (unpublished) at 9 kMc on Various Ferrites
H∥ Slab	1100	0·51	0·42
	1600	0·55	0·50
	2900	0·69	0·65
H⊥ Slab	1600	0·65	0·62

Here we shall only give a qualitative outline of the minimization procedure, which is largely a matter of numerical computation. We assume η_k to be independent of k. It is clear first of all that θ_k must neither be zero nor $\eta/2$, otherwise $\omega_{s\,crit}$ would be infinite. Hence the spin waves of interest propagate in some oblique direction. Furthermore, as is to be expected from symmetry, $\omega_{s\,crit}$ is independent of φ_k the longitude of $\boldsymbol{k}$. Hence the $\boldsymbol{k}$ vector of a critical spin wave can lie anywhere on a circular cone about the magnetization direction. Let us now proceed to minimize (26), subject to (27), at various fixed values of ω_H. At sufficiently small ω_H, the equation $\omega/2 = \omega_k$ will always have a real solution for $|\boldsymbol{k}|$ in terms of θ_k and ω_H. Hence at sufficiently small ω_H, $\omega_{s\,crit}$ may be expressed in terms of θ, and ω_H. In fact the equation $\omega_k = \omega/2$ gives

$$\omega_H - \omega_M N_z + \omega_{ex} l^2 k^2 = -\frac{\omega_M \sin^2\theta_k}{2} + \\ + \tfrac{1}{2}\sqrt{(\omega_M{}^2 \sin^4\theta_k + \omega^2)} \tag{28}$$

so that (26) becomes (neglecting η_0)

$$\omega_{s\,crit} = \\ \frac{4\omega(\omega_{res} - \omega)\eta_k}{\omega_M \sin\theta_k \cos\theta_k [\omega - \omega_M \sin^2\theta_k + \sqrt{(\omega_H{}^2 \sin^4\theta_k + \omega^2)}]} \tag{29}$$

This expression, minimized with respect to θ_k, varies

with respect to ω_H simply as $\omega_{res}-\omega$ varies with ω_H. For example in a sphere $\omega_{res}=\omega_H$, so that the minimum critical signal declines linearly as ω_H is increased towards resonance. However, (19) applies only so long as ω_H is sufficiently small, so that (28) gives a positive value for k^2 when θ_k has the value minimizing (29). Once ω_H is increased beyond the value for which k^2 is zero according to (28), the further course of $\omega_{s\,\mathrm{crit}}$ (min.) is determined by solving (28) for θ_k in terms of ω_H, and equating k to zero there and in (26). Then

$$\sin^2\theta_k=\frac{\omega^2-4(\omega_H-\omega_M N_z)^2}{4\omega_M(\omega_H-\omega_M N_z)} \tag{30}$$

and

$$\omega_{s\,\mathrm{crit}}(\mathrm{min})=\frac{2\omega\eta_k(\omega_{\mathrm{res}}-\omega)}{\omega_M\sin\theta_k\cos\theta_k[(\omega/2)+\omega_H-\omega_M N_z]}. \tag{31}$$

No further minimizing is now required. In most experiments, $\omega_{s\,\mathrm{crit}}$ (min) will continue to decrease as ω_H increases beyond the point at which $k=0$. Eventually $\omega_{s\,\mathrm{crit}}$ (min) passes through an absolute minimum some hundreds of oersteds below resonance, and then increases again, towards infinity as ω_H approaches $\omega/2+\omega_M N_z$ (i.e. $\theta_k\to 0$). Thus the subsidiary absorption is always confined to a range of fields such that

$$\omega_H<\omega/2+\omega_M N_z \tag{32}$$

and in the more common experiments (on nickel or manganese ferrite spheres ($4\pi M$ 3300 oe) at 10 kMc or higher frequencies) this upper limit is well below resonance.

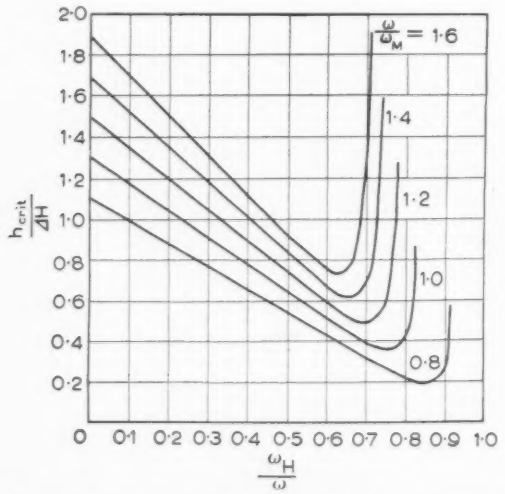


FIG. 1. Threshold signal for the instability connected with the subsidiary absorption in the case of a sphere, as a function of applied d.c. field, for various values of saturation magnetization. (The Landau-Lifshitz form of the damping term has been used in the computation.)

It is to be assumed that much of the subsidiary absorption involves spin waves with $|\boldsymbol{k}|\sim 0$, and as noted earlier, the demagnetizing field used in the analysis is then incorrect. That one still achieves good qualitative and even quantitative agreement with experiment is probably due to the fact that the frequencies of WALKER'S modes all fall between the upper and lower limits of ω_k for $|\boldsymbol{k}|=0$, which are attained at $\theta=\pi/2$ and $\omega=0$ respectively, and that they become increasingly dense as $\theta\to 0$.

Hence there is never any shortage of modes satisfying the subharmonic condition (27) even though the description in terms of plane waves is not correct (see also section 5).

4. MERGING OF THE SUBSIDIARY ABSORPTION WITH THE MAIN RESONANCE

Case (b), while it requires special experimental arrangements, is extremely interesting, since it involves an exceptionally small threshold signal, of order η^2/ω_M. (For example with $4\pi M\sim 3000$ oe, $\eta/\gamma\sim$ ten oersteds, the threshold signal is about 1/30 r.f. oersted.) In order to satisfy $\omega_k=\omega/2$, $\omega_{\mathrm{res}}=\omega$ simultaneously, we must have $\omega_H=\omega+(N_z-N_T)\omega_M$ and

$$(\omega-N_T\omega_M+\omega_{\mathrm{ex}}l^2k^2)\times$$

$$\times(\omega-N_T\omega_M+\omega_{\mathrm{ex}}k^2+\omega_M\sin\theta_k)=\omega^2/4.$$

Provided

$$N+\omega_M>\omega/2 \tag{33}$$

the last equation can be satisfied by a real $\boldsymbol{k}$. When (33) is satisfied by an adequate margin, there is no doubt that the minimum of (26) occurs very close to $\omega_{\mathrm{res}}=\omega$, $\omega/2=\omega_k$, since the radicand has a very deep minimum there. The minimum threshold is then exactly at resonance, its precise value depends on whether ω_H at resonance is such that (28) gives a real value of k when θ_k minimizes

$$\omega_{s\,\mathrm{crit}}=\frac{2\omega_k\eta_k\eta_0}{\omega_M\sin\theta_k\cos\theta_k[(\omega/2)+\omega_H-\omega_M N_2+\omega_{\mathrm{ex}}l^2k^2]}$$

or whether ω_M is such that k must be zero, in accordance with the remarks in section 3. The latter case is the more common, and then the threshold is given by

VOL. 1 1956/57

$$\omega_{s\,\mathrm{crit}}=\frac{2\omega\gamma_k\gamma_0}{\omega_M\sin\theta_k\cos\theta_k[(\omega/2)+\omega-N_T\omega_M]}$$

with θ_k given by (27), subject to $\omega_H=\omega+(N_z-N_T)\omega_M$.

Since no such experiment had been reported in the literature, one was conducted by the author on a thin, high-resistivity, single-crystal manganese-ferrite disc, and on a similar disc of low-resistivity nickel ferrite. The discs were magnetized in their own plane, so that the geometries were not spheroidal about the direction of the d.c. field. This has no effect on the spin-wave analysis, at least to second order in the spin-wave amplitudes. The expression for ω_k is likewise unchanged, since it involves only N_z. The only appreciable change will occur in the uniform precession which now is no longer circularly polarized. The threshold field will therefore be different by at most a factor two. The condition that $2\omega_k=\omega=\omega_{\mathrm{res}}=\sqrt{[\omega_H(\omega_H+\omega_M)]}$ should hold simultaneously now becomes

$$\omega_M>\frac{3}{2}\omega,$$

a condition which will hold in nickel and manganese ferrite up to about 7 kMc. A standard resonance experiment showed that an anomalous change (to be discussed in section 6) occurred in the main resonance at a signal field of only about one hundredth of an oersted for the manganese-ferrite sample, and about 10 db above this value for the nickel ferrite.

5. PREMATURE SATURATION OF THE MAIN RESONANCE, AND THE DEGENERATE SPIN-WAVE SPECTRUM

We have shown that the subsidiary absorption is related to instability arising from double terms in the equation of motion for the spin waves. Similarly the premature saturation of the main resonance can be traced to an instability arising from triple terms.

Returning to equation (22), restricting it to spin waves for which $\omega_k\sim\omega$, and neglecting all terms with rapidly varying coefficients, we find that all double terms disappear, leaving only terms of second order in a_0°,

$$\dot{b}_k^\circ=i\frac{\lambda_k^2}{2}\left(\omega_{ex}l^2k^2-\omega_M N_T+\omega_M\frac{k_z^2}{k^2}\right)\times$$
$$\times a_0^{\circ 2}b_{-k}^{\circ *}e^{2i(\omega-\omega_k)t}.$$

Here the term involving b_k°, which only produces a slight frequency shift has been neglected. A more detailed calculation, taking it into account produces no change in the result for the threshold signal, and merely changes the wave number of the critical spin wave slightly. Writing

$$\frac{\lambda_k^2}{2}\left(\omega_{ex}l^2k^2-\omega_M N_T+\omega_M\frac{k_z^2}{k^2}\right)=\xi_k$$

and

$$b_k^\circ=v_ke^{i(\omega-\omega_k)t}$$

we find

$$[d/dt+i(\omega-\omega_k)]v_k=i\xi_k(a_0^\circ)^2v_{-k}^*.$$

Elimination of v_{-k}^* from this, and the "adjoint" equation gives

$$[d^2/dt^2+(\omega-\omega_k)^2-\xi_k^2|a_0^\circ|^4]v_k=0.$$

Hence $b_k=b_k^\circ e^{(i\omega_k-\eta_k)t}$ will increase exponentially with time, if

$$\xi_k^2|a_0^\circ|^4>(\omega-\omega_k)^2+\eta_k^2$$

or, in terms of the signal, if $\omega_s>\omega_{s\,\mathrm{crit}}$, where

$$\omega_{s\,\mathrm{crit}}=\frac{[(\omega-\omega_k)^2+\eta_k^2]^{1/4}[(\omega-\omega_{\mathrm{res}})^2+\eta_0^2]^{1/2}}{\xi_k^{1/2}}.$$

In general, this threshold is of order ω, that is, much too large for practical purposes unless

$$\omega_k=\omega.$$

Even then it is still too large, of order $\omega\sqrt{\eta_k/\omega_M}$), unless we also have

$$\omega_{\mathrm{res}}=\omega.$$

If k is such that

$$\omega_{\mathrm{res}}=\omega=\omega_k \tag{35}$$

the threshold will be

$$\omega_{s\,\mathrm{crit}}=\eta_0\frac{\eta_k^{\frac{1}{2}}}{\xi_k^{\frac{1}{2}}} \tag{36}$$

which is of order $\eta\sqrt{(\eta/\omega_M)}$ and is within reach

experimentally. That the equation $\omega_k = \omega_{res}$ can be satisfied for all practical geometries by a real manifold of $\boldsymbol{k}$ values is a relatively new result. Our formula for ω_k (plotted in Fig. 2)

$$\omega_k^2 = (\omega_H - \omega_M N_z + \omega_{ex} l^2 k^2) \times$$
$$\times (\omega_H - \omega_M N_z + \omega_{ex} l^2 k^2 + \omega_M \sin^2\theta_k)$$

differs from the conventional one:

$$\omega_k^2 = (\omega_H + \omega_{ex} l^2 k^2) \times$$
$$\times (\omega_H + \omega_{ex} l^2 k^2 + \omega_M \sin^2\theta_k).$$

The reason for the difference is due to the fact that the conventional formula applies to the unbounded medium. In any real sample the spin waves move, not in the applied d.c. field H_0, but in a field $H_0 - 4\pi M_0 N_z$, demagnetized by the pole distribution due to the constant part of M_0. On the basis of the conventional spectrum, only prolate spheroids about H_0 were capable of the degeneracy $\omega_k = \omega_{res}$ for a real manifold of spin waves. For the actual spectrum the degeneracy exists for all geometries except for the infinitely thin normally magnetized, disc, whose resonant frequency $\omega_H - \omega_M$, lies at the lowest point $k = 0$, $\theta = 0$ of the spin wave spectrum, so that the degenerate manifold has shrunk to a point. In the present case the degeneracy has the consequence that power can be transferred from the uniform precession at resonance to the degenerate manifold of spin waves in such a way as to lead to instability at sufficient signal levels. However, even at low signal levels, other coupling mechanisms, such as nonuniformities in the lattice, etc., may transfer power from the uniform precession. In this way a finite, low-level line width may be deduced by means of a Weisskopf-Wigner type of calculation utilizing the degeneracy. Such a calculation is reported in reference 6.

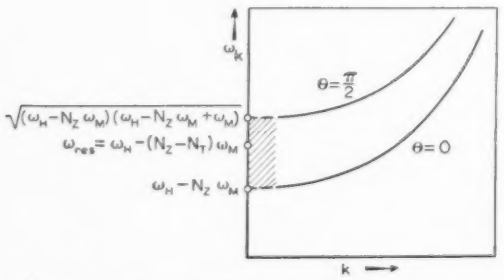


FIG. 2. The spin-wave spectrum, showing the degeneracy of the uniform precession frequency with that of a continuum of spin-wave modes. In the cross-hatched region, plane spin waves are not strictly correct modes of the system.

Proceeding with the present problem, the threshold 36 may be minimized with respect to k subject to $\omega_k = \omega_{res}$. The result is that a z-directed spin wave has the lowest threshold. The corresponding k is found by writing $\theta_k = 0$ in $\omega_k = \omega$, and is given by

$$\omega_{ex} l^2 k^2 = \omega_M N_T.$$

This result applies both in the case in which ηk is assumed independent of k, and in the case in which the damping has Landau-Lifshitz form. The least value of the threshold is

$$\omega_{s\,crit} = \eta_0 \sqrt{(\eta_k/\omega_M)} \qquad (37)$$

For Landau-Lifshitz damping $\eta_k = \alpha A_k = \alpha\omega_k = \alpha\omega = \eta_0$, for $z-$ directed spin waves, and the threshold is then $\eta_0 \sqrt{(\eta_0/\omega_M)}$.

Table 2 shows the agreement between (37) and the B.W.-D results for the onset of apparent

Table 2. Computed and measured threshold signals on the main resonance (when it is distinct from the subsidiary resonance). For definiteness it has been assumed that $\Delta H_k = \Delta H = 1/\gamma T_1$. Agreement is actually not as good as it seems, since the theory assumes circular polarization of the signal, while the measurement referred to linear polarization. The resulting discrepancy is probably due to the uncertainty in ΔH_k, ΔH.

$\Delta H = \frac{1}{\gamma T_1}$ (oe)	r.f. threshold field (oe)		
	Theory	Experimental	
		Bloembergen and Wang	Damon "Sample 7"
51	8·0	9·0	
46	7·0	5·5	
36	5·0	4·0	
23	2·5	1·0	≈ 2·0
7	0·45	0·55	

VOL. 1 1956/57

saturation. For order of magnitude comparison η_k and η_0 were both taken to equal $1/T_1$ where T_1 was the relaxation time measured by noting the change in the spatial mean value of M_z with applied power. T_1 was used in preference to the observed linewidth.* The latter may be unduly widened by macroscopic irregularities, whereas T_1, derived from the actually absorbed power, presumably measured the relaxation time of a small regular region. Since the break in the susceptibility curve will begin when the first few small regions in the sample go unstable, it is T_1 that is most likely to lead to the correct threshold value. The agreement may not be as good as it would seem from the table, since the theory assumes a circularly polarized signal field, while the measurement apparently referred to linear polarization. The resultant discrepancy of a factor two might well be due to our inability to estimate the present uncertainty in η_k, η_0.

6. THE STEADY STATE BEYOND THE THRESHOLD SIGNAL

All that has been established so far is that the observed onset of anomalous absorption effects coincides approximately with the theoretical onset of unstable growth of certain spin waves. This growth will, of course, not continue indefinitely, and in order to predict observations beyond the threshold, it is necessary to determine in what new magnetic state the sample will settle down. It is at this point that the theory becomes nonlinear in a less trivial sense.

We might imagine an experiment in which the uniform precession angle is very rapidly opened out to a value beyond critical by the rising edge of a large microwave pulse. The resulting exponential increase in the relevant spin-wave amplitudes will continue almost unabated until they are large enough to appreciably modify the uniform precession through the appropriate nonlinear terms in equation (12). It will turn out that the primary effect of these growing non-linear terms is to progressively enhance the losses of the uniform precession, causing the precession angle to diminish, although the signal field remains unchanged. Consequently the growth rates of the relevant spin waves are reduced, but they still continue to increase, causing a further drop in the precession angle. Finally, this angle settles down at a value just below critical, the growth of the spin waves ceases, but their amplitudes have attained large nonthermal values.

The exact calculation of the complete transient ending in this final state is difficult; the final state itself, on the other hand, can be deduced by analyzing an idealized experiment in which the signal level is increased from zero to its final value above the threshold in a series of infinitesimal steps. After each step the system is permitted to attain its steady state. When the signal exceeds the threshold, only one steady state is found, and it is therefore reasonable to expect that it is the one following the transient in the more usual experiment previously described. Our first task is to examine the variation of the spin-wave level in the course of the "adiabatic" experiment. Far below the signal threshold, the spin-wave level is maintained solely by thermal agitation. The precise mechanism of this agitation, which ultimately must derive from the coupling of the spin system to the lattice, is of no consequence in the present problem. We may simply describe the agitation as a random magnetic driving field, which is just sufficiently large to maintain the spin-wave level at a value appropriate to the Bose-Einstein statistics, in the face of the phenomenological losses described by η. The actual magnitude is never required, however, for the thermal field finally turns out to be a purely auxiliary quantity. We denote it by $\boldsymbol{H}^{\text{th}}(t, \boldsymbol{r})$, a random function of t and $\boldsymbol{r}$, and begin by expanding it in a spatial Fourier series

$$\boldsymbol{H}^{\text{th}}(t, \boldsymbol{r}) = \Sigma\, \boldsymbol{H}_k{}^{\text{th}}(t) e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{38}$$

whose coefficients, $\boldsymbol{H}_k{}^{\text{th}}$, are random functions of time. For convenience we may regard $\boldsymbol{H}_k{}^{\text{th}}$, or more accurately, the quantity

$$\nu_k(t) = \gamma(H_{kx}{}^{\text{th}} + iH_{ky}{}^{\text{th}}) \tag{39}$$

as the driving field of the normal modes b_k, rather than of the undiagonalized a_k's. When the coupling to the uniform precession is neglected, the motion of b_k is given by

$$\dot{b}_k = i\Omega_k b_k + \nu_k(t). \tag{40}$$

* T_1 was varied in the B.W. experiment by varying the temperature.

When the coupling to a_0 is included, and b_k taken in the form $b_k^\circ(t)e^{i\omega_k t}$, we find, restricting the analysis to the $\omega = 2\omega_k$-type of instability,

$$\dot{b}_k^\circ = \rho_k a_0^\circ b_{-k}^{\circ *} e^{i(\omega - 2\omega_k)t} + \\ + \nu_k(t) e^{-i\omega_k t + \eta_k t}. \tag{41}$$

We have assumed here that, in the coupling term, a_0° may be regarded as the time average of the amplitude of a_0. In actual fact a_0° also has thermally fluctuating components, especially beyond the threshold. Writing

$$b_k^\circ = s_k e^{i[(\omega/2) - \omega_k]t + \eta_k t}$$

(41) becomes

$$\dot{s}_k + \left[i\left(\frac{\omega}{2} - \omega_k\right) + \eta_k \right] s_k \\ = \rho_k a_0^\circ s_{-k}^* + \nu_k(t) e^{-i(\omega/2)t} \tag{42}$$

In solving this equation we shall follow the procedure of UHLENBECK and WANG[14] for analysing random processes. We assume that $\nu_k(t) = 0$ outside a large but finite time interval $(-T, +T)$, and assume that $\nu_k(t)$ has a Fourier transform:

$$\tilde{\nu}_k(\xi, T) = \frac{1}{\sqrt{(2\pi)}} \int_{-T}^{+T} \nu_k(t) e^{-i\xi t}\, dt$$

which is such that $\underset{T \to \infty}{\text{Lt}} \frac{1}{2T} |\tilde{\nu}_k(\xi, T)|^2$ exists. Thus we have

$$\nu_k(t) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{+\infty} \tilde{\nu}_k(\xi, T) e^{i\xi t}\, d\xi.$$

Substituting this expression for $\nu_k(t)$ in (42) and solving it and the "adjoint" equation, we find:

$$s_k(t) = \frac{\rho_k a_0^\circ}{\sqrt{(2\pi)}} \int \frac{\nu_k^*(\xi) e^{-i[\xi - (\omega/2)]t}\, d\xi}{\left(\frac{\omega}{2} - \omega_k\right)^2 - \left(\frac{\omega}{2} - \xi\right)^2 + \eta_k^2 - \rho_k \rho_k^* |a_0^\circ|^2 - 2i\eta_k \left(\xi - \frac{\omega}{2}\right)} + \\ + \frac{1}{\sqrt{(2\pi)}} \int \frac{[\eta_k + i(\omega_k + \xi - \omega)] \nu_k(\xi) e^{-i[\xi - (\omega/2)]t}\, d\xi}{\left(\frac{\omega}{2} - \omega_k\right)^2 - \left(\frac{\omega}{2} - \xi\right)^2 + \eta_k^2 - \rho_k \rho_k^* |a_0^\circ|^2 + 2i\eta_k \left(\xi - \frac{\omega}{2}\right)} \\ \equiv \frac{1}{\sqrt{(2\pi)}} \int P_k(\xi) e^{-[i\xi - (\omega/2)]t} + Q_k(\xi) e^{i[\xi - (\omega/2)]t}\, d\xi, \text{ say.} \tag{43}$$

This result which determines the spin-wave excitation in the presence of the uniform precession, must be substituted in the double terms of equation (12). Higher terms may be neglected in this approximation. We attempt a solution of the form

$$a_0 = f(t) e^{i\omega t}$$

where $f(t)$ is a bounded random function of time with a finite mean value equal to a_0°. Then (12) becomes, to second order, and with the damping term included:

$$-i\dot{f} + [\omega - \omega_{\text{res}} - i\eta_0] f \\ = \omega_s - \frac{\omega M}{2} e^{-i\omega t} \sum_{k \neq 0} a_{-k} (a_k k_T^* + a_k k_T) \frac{k_z}{k^2}.$$

Averaging this expression over $(-T, +T)$ we obtain, since f is bounded,

$$[\omega - \omega_{\text{res}} - i\eta_0] a_0^\circ = \omega_s - \frac{\omega M}{2} \times \\ \times \frac{1}{2T} \int_{-T.}^{+T} dt\, e^{-i\omega t} \sum_{k \neq 0} a_{-k} (a_k k_T^* + a_{-k}^* k_T) \frac{k_z^2}{k^2}. \tag{44}$$

We have

$$a_k = \lambda_k b_k - \mu_k b_{-k}^* \\ = \lambda_k s_k e^{i(\omega/2)t} - \mu_k s_{-k}^* e^{-i(\omega/2)t}. \tag{45}$$

In substituting 45 in the integrand in 44, we may make use of the fact that since the agitation is random in space as well as in time, the phase of ν_k is a rapidly fluctuating function of k; hence in sums of the form $\Sigma F(k,k') \nu_k \nu_{k'}^*$ only the terms with $k' = k$ need be retained, the others averaging to zero. Sums of the form $\Sigma F(k,k') \nu_k \nu_{k'}$ vanish altogether. Hence we have, for purposes of summation over k:

VOL. 1 1956/5

$$a_k a_{-k} = \lambda_k{}^2 s_k s_{-k} e^{i\omega t} - \mu_k{}^2 s_{-k}{}^* s_k{}^* e^{-i\omega t} - \lambda_k \mu_k (s_k s_k{}^* + s_{-k} s_{-k}{}^*)$$

$$= \lambda_k{}^2 \frac{e^{i\omega t}}{2\pi} \iint d\xi d\xi' [P_k(\xi) Q_{-k}(\xi') + Q_k(\xi') P_{-k}(\xi)] e^{i(\xi'-\xi)t}$$

$$\frac{-\mu_k{}^2 e^{-i\omega t}}{2\pi} \iint d\xi d\xi' [P_k{}^*(\xi) Q_{-k}{}^*(\xi') + Q_k{}^*(\xi') P_{-k}{}^*(\xi)] e^{i(\xi'-\xi)t}$$

$$\frac{-\lambda_k \mu_k}{2\pi} \iint d\xi d\xi' [P_k(\xi) P_k{}^*(\xi') + Q_k(\xi') Q_k{}^*(\xi) + P_{-k}(\xi) P_{-k}(\xi') \times Q_{-k}{}^*(\xi') Q_{-k}{}^*(\xi)] e^{i(\xi'-\xi)t}. \quad (46)$$

Of these terms only the first survives substitution in equation (44) and integration over t, since the time integral, if performed first, has the form

$$\int_{-T}^{+T} e^{i(\xi'-\xi-\omega)t} e^{i\omega t} dt = \frac{2\sin(\xi-\xi')T}{(\xi-\xi')}$$

which for sufficiently large T may be identified with 2π times the delta function $\delta(\xi-\xi')$.

Integration over ξ' then leaves the integrand proportional to $(1/2T)|\nu_k(\xi,T)|^2$ (through the term $P_k Q_{-k}$), etc. As T tends to infinity this expression has a finite limit. On the other hand integration involving the remaining terms in (46) produces $\delta(\xi-\xi'-2\omega)$ and $\delta(\xi-\xi'-\omega)$. Integration over ξ' then leads to terms such as $(1/2T)$ $[\nu_k(\xi,T)\nu_k{}^*(\xi-2\omega,T)]$ which may be assumed to vanish as $T\to\infty$, if the spectral components are uncorrelated. Thus the first term in (46), when substituted in (44) leads to

$$\frac{-\omega M}{2} \sum_k \lambda_k{}^2 k_T{}^* \int \frac{1}{2T} [P_k(\xi) Q_{-k}(\xi) + Q_k(\xi) P_{-k}(\xi)]\, d\xi$$

while all other terms give zero. In the same way, the only term in $a_{-k}a_{-k}{}^*$ surviving integration and summation has the form $\lambda_k \mu_k s_{-k} s_k e^{i\omega t}$. Combining all these results and recalling that

$$\rho_k = i\omega_M \lambda_k (\mu_k k_T{}^* - \lambda_k k_T) \frac{k_z}{k^2}$$

we find that (44) becomes

$$(\omega - \omega_{\rm res} - i\eta_0) a_0{}^\circ = \omega_s + i a_0{}^\circ \int \sum_k \frac{\rho_k \rho_k{}^* [\eta_k + i(\omega_k + \xi - \omega)] T_k(\xi)\, d\xi}{\left[\left(\frac{\omega}{2} - \omega_k\right)^2 - \left(\frac{\omega}{2} - \xi\right)^2 + \eta_k{}^2 - \rho_k \rho_k{}^* |a_0{}^\circ|^2\right]^2 + 4\eta_k{}^2 \left(\xi - \frac{\omega}{2}\right)^2} \quad (47)$$

where

$$T_k = \frac{1}{2} \operatorname*{Lt}_{T\to\infty} \frac{1}{2T} [|\nu_k(\xi,T)|^2 + |\nu_{-k}(\xi,T)|^2].$$

Rearranging terms and taking squares of absolute values, we find

$$|a_0{}^\circ|^2 = \frac{\omega_s{}^2}{(\omega-\omega_{\rm res})^2 + (\eta_0 + S)^2} \quad (48)$$

where S stands for the integral in (47), and still depends on $|a_0{}^\circ|^2$. The solution of (48) is traced by graphical means as follows: The left-hand side of (48), plotted versus $|a_0{}^\circ|^2$, is always a straight line with unit slope (see Fig. 3). The right-hand side, for small $|a_0{}^\circ|^2$, is practically independent of $|a_0{}^\circ|^2$ and equal to $\omega_s{}^2/[(\omega-\omega_{\rm res})^2+\eta_0{}^2)]$, since S is then merely a small thermal contribution to the loss, which, as experience with most ferrites shows, can be neglected.† The right-hand side of (48) plotted against $|a_0{}^\circ|^2$ is thus an almost horizontal straight line, until $|a_0{}^\circ|^2$ approaches to within a small neighborhood of the least value of the expression

$$\frac{\eta_k{}^2 + \left(\frac{\omega}{2} - \omega_k\right)^2}{\rho_k \rho_k{}^*}$$

† If we write $a_0{}^\circ = 0$ in the integral S, we have what is in effect the contribution to the loss at finite temperatures, calculated by KEFFER[11] from the zero-point motion at absolute zero. However, S is appreciable only if $\omega = 2\omega_k$ can be satisfied for at least some of the terms on the summation, and, except in the special arrangement discussed in section 4, this is possible only some hundreds of oersteds below resonance.

considered as a function of k. That least value is, of course, the threshold precession discussed in section 3 (equation 25). From then on S rises very sharply to infinity at $|a_0^\circ|^2 = |a_0^\circ|^2_{\text{crit}}$, since one term in the integrand then approaches a second order pole at $\xi = \omega/2$. Furthermore, provided we regard the $\boldsymbol{k}$ vectors as forming a continuum, S remains infinite as $|a_0^\circ|^2$ is increased beyond the threshold, since there will then always be *some* term in the integrand which goes to infinity. If the $\boldsymbol{k}$ vectors are not a continuum, S will go through a rapid succession of infinities, but as is seen by inspection, S will remain very large compared with ω, even at the intervening minima, if the spacing between successive k vectors is not too great. Estimates of the spacing can of course be made, but their value is questionable since the modes involved are frequently not true spin waves, and their density would have to be evaluated on a different basis. We therefore make the essumption that S remains very large once $|a_0^\circ|^2$ exceeds the critical value. Consequently the right-hand side of (48) as a function of $|a_0^\circ|^2$, is an almost horizontal straight line, until $|a_0^\circ|^2$ approaches to within a very small neighborhood of the critical value,* then plunges to zero and remains at that value for all practical purposes.

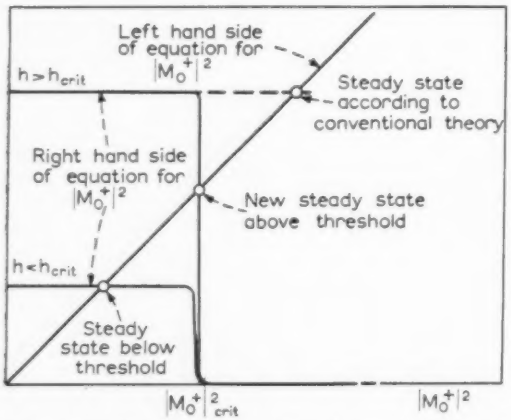

Left hand side of equation for $|M_0^+|^2$
$h > h_{crit}$
Steady state according to conventional theory
Right hand side of equation for $|M_0^+|^2$
New steady state above threshold
$h < h_{crit}$
Steady state below threshold
$|M_0^+|^2_{crit}$
$|M_0^+|^2$

FIG. 3. Graphical solutions of equation (48), for signal fields below and above the threshold for the subsidiary absorption.

The solution of (48) is the intersection of the left- and right-hand sides (see Fig. 3). So long as ω_s is small the intersection is on the horizontal part of the right-hand side of (48), and, of course, represents the conventional Kittel solution of the problem. As ω_s is increased, the horizontal part of the curve rises parallel to itself, and the intersection moves towards the breaking point. Finally as ω_s reaches the least value of $\omega_{s\,\text{crit}}$ (at fixed ω_H), the solution is exactly $|a_0^\circ|^2_{\text{crit}}$. For $\omega_s > \omega_{s\,\text{crit}}$, the intersection lies on the very rapidly declining part of the curve, and from then on $|a_0^\circ|^2$ remains very close to, but just below $|a_0^\circ|^2_{\text{crit}}$ no matter how large ω_s.

The calculation of the observed imaginary part χ'' of the susceptibility in cases a and b of section 3 now presents no further difficulty. When $\omega_s > \omega_{s\,\text{crit}}$, we can write down two expressions for $|a_0^\circ|^2_{\text{crit}}$:

$$|a_0^\circ|_{\text{crit}}^2 = \frac{\omega_s{}^2{}_{\text{crit}}}{(\omega-\omega_{\text{res}})^2+\eta_0{}^2} = \frac{\omega_s{}^2}{(\omega-\omega_{\text{res}})^2+(\eta_0+S)^2}$$

the first of which is true by definition, and the second of which is true with only negligible error. Thus we can solve for the effective loss parameter:

$$(\eta_0+S) = \sqrt{\left(\frac{P}{P_{\text{crit}}}-1\right)(\omega-\omega_{\text{res}})^2+\eta_0{}^2\frac{P}{P_{\text{crit}}}}, \qquad (P > P_{\text{crit}})$$

where P/P_{crit} is the power ratio $\omega_s{}^2/\omega_s{}^2{}_{\text{crit}}$.

χ'', which is proportional to the power dissipated per unit applied power, i.e., to $(\text{Im}(a_0^\circ/\omega_s)$, is therefore

$$\chi''_{P>P_{\text{crit}}} = \chi_0{}''\frac{P_{\text{crit}}}{P}\times \times\sqrt{\left(\frac{P}{P_{\text{crit}}}-1\right)\frac{(\omega-\omega_{\text{res}})^2}{\eta_0{}^2}+\frac{P}{P_{\text{crit}}}} \qquad (49)$$

where $\chi_0{}''$ is the ordinary small signal susceptibility.

When, as is more usual, the subsidiary absorption occurs well below the main line the second term in the radicand of (49) can be neglected by comparison with the first as soon as P exceeds P_{crit} only slightly. We then have, approximately

$$\frac{\chi''_{P>P_{\text{crit}}}}{\chi_0{}''} = \frac{\omega-\omega_{\text{res}}}{\eta_0}\times \times\frac{P_{\text{crit}}}{P}\sqrt{\left(\frac{P}{P_{\text{crit}}}-1\right)}$$

* Of order $\sqrt{\left(\frac{\eta_k T_k}{\eta_0 \rho_k \rho_k{}^*}\right)}$.

which rises from zero at $P = P_{crit}$ (strictly, from unity, according to eq. (49) to a maximum $(\omega - \omega_{res}/2\eta_0)$ at $P = 2P_{crit}$, and then declines steadily to zero as $P \to \infty$. Fig. 4 shows the comparison of the theory with unpublished measurements of Dr. H. E. D. SCOVIL on a manganese-ferrite sphere at X band. Agreement is quite good except on the high-power side of the maximum. There Dr. SCOVIL noted the onset of further as yet unexplained instabilities, which added to the susceptibility.

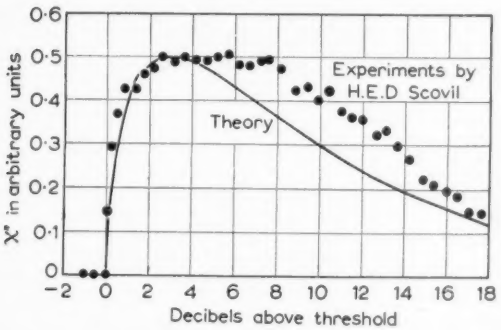


FIG. 4. Susceptibility versus power above threshold at the subsidiary absorption peak for a sphere of manganese ferrite.

The B.W.-D data, too, are at least in good qualitative agreement with the theory. These authors did not measure χ'' at the subsidiary resonance, but they did give the changes in the observed Z-component of the magnetization upon application of the signal pulse. As is shown in the appendix, that change will be proportional to $\chi''P$, provided P is sufficiently in excess of P_{crit} at the subsidiary absorption. From (49) it follows that $(M\text{-}M_z)^2$ should therefore vary approximately as $[(P/P_{crit})-1]$. Fig. 5 shows that such is very nearly the case for the high-resistivity nickel-ferrite specimen of B.W.-D. It is interesting to note that, apparently, DAMON's Sample ("No. 7") had a lower threshold than that of B.W.

Some further confirmation of the theory was obtained in the author's measurement in which subsidiary and main resonance merge. In that case, $\omega = \omega_{res}$, and eq. (49) gives

$$\frac{\chi_P'' > P\text{crit}''}{\chi_0''} = \sqrt{\frac{P_{crit}}{P}}$$

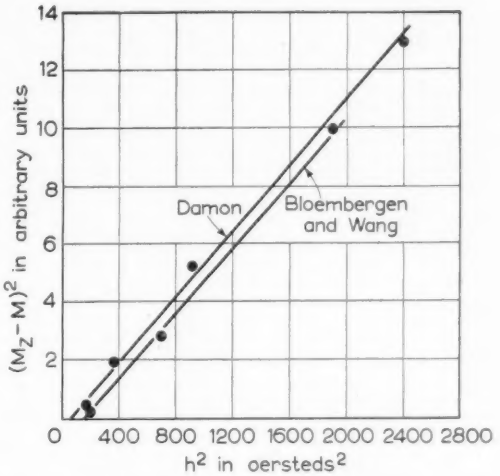


FIG. 5. $(\overline{M}_z - M)^2$ as a function of signal power at the subsidiary absorption in the B.W.-D experiments. The theory predicts a straight line intercepting the abscissa at the critical power. The material was stoichiometric (high-resistivity) nickel ferrite.

Fig. 6 shows the experimental results for a high-resistivity manganese disc and a low-resistivity nickel-ferrite disc, magnetized on their own planes. The manganese-ferrite disc not only had a very low threshold, but also followed the $(1/\sqrt{P})$-law quite closely. The curve for the non-stoichiometric nickel ferrite declined much too slowly, which might be ascribed to nonuniformities in the material.

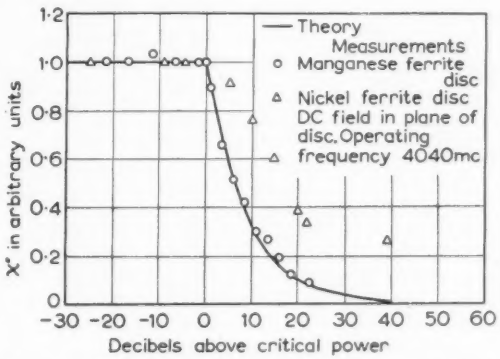


FIG. 6. χ'' versus power when subsidiary and main resonance coincide. The manganese-ferrite disc is in close accord with theory. The nickel–ferrite sample (nonstoichiometric) is in poor agreement. Also the value of its threshold exceeded that of manganese ferrite by 10 db.

7. FURTHER INSTABILITIES AND SOME UNSOLVED PROBLEMS

The present theory still suffers from a number of shortcomings. For example in deriving the behavior beyond the threshold we assumed that a_0 in equation (41) could be replaced by $a_0^\circ e^{i\omega t}$ and a_0° be treated as a constant. This will be nearly true for a moderate range of power above critical, but eventually the double terms in (12), not all of which have the "right" frequency, will modulate a_0 rapidly, and at frequencies such that a solution of type (43) would be seriously upset. A further shortcoming of the theory of the final state above the threshold is that its extension to the instability at the main resonance (when it is distinct from the subsidiary peak) is very difficult. The reason is that the third-order terms involve not only $a_0^2 a_k$, but also terms such as $a_k \Sigma a_k a_k^*$ etc., which, collectively, became as large $a_0^2 a_k$ beyond the threshold. These added terms complicate the algebra to a great, perhaps impossible, extent.

Another unresolved question is that of the great variety of further subsidiary absorptions noted by Dr. H. E. D. SCOVIL. These are much less pronounced than those of B.W.-D, and occur at a higher threshold, but they appear to be spread from very low d.c. fields right up to the major subsidiary absorption. It is not impossible that these further absorptions correspond to unstable growth of further spin waves that feed upon those already sustained by the uniform precession. On the other hand, they may be connected with the "true" modes that should replace the spin waves of long wavelength.

Acknowledgements—The author is indebted to Dr. P. W. ANDERSON, A. M. CLOGSTON, and L. R. WALKER for a great many helpful discussions.

Thanks are also due to Drs. H. E. D. SCOVIL and M. T. WEISS for permission to quote their unpublished data.

APPENDIX

Relation between the average of $M-\overline{M}_z$ and the power dissipated in the sample.

It follows from the definition of χ'' that the dissipated power is proportional to $\chi''P$, where P is the applied power. On the other hand, the energy stored in the oscillators b_k and a_0 is proportional to

$$\omega|a_0|^2+\Sigma\,\omega_k|b_k|^2$$

and the rate of dissipation is proportional to

$$\omega\eta_0|a_0|^2+\Sigma\,\omega_k\eta_k|b_k|^2.$$

For the subsidiary absorption, only spin waves for which $\omega_k=\omega/2$, and whose $\boldsymbol{k}$-vectors form a thin conical shell about the magnetization direction, ever become important. Assuming, as is very likely, that η_k is independent of $\boldsymbol{k}$ on this cone and recalling that above the threshold, $|a_0^\circ|$ 'sticks' at its critical value, the rate of dissipation is seen to vary as

$$\omega[\eta_0|a_0^\circ|^2_{\text{crit}}+\tfrac{1}{2}\eta_k\,\Sigma'|b_k|^2] \qquad \text{(i)}$$

where the sum Σ' extends only over spin waves on the critical cone. On the other hand the space and time average of $1-m_z$ is

$$\begin{aligned}\langle\delta m_z\rangle &=\tfrac{1}{2}\langle\,\Sigma\,a_k a_k^*\rangle\\ &=\tfrac{1}{2}|a_0^\circ|^2_{\text{crit}}+\tfrac{1}{2}\,\Sigma'|b_k|^2\cosh\psi_k\\ &=\tfrac{1}{2}[|a_0^\circ|^2_{\text{crit}}+\cosh\psi_k\,\Sigma'|b_k|^2] \qquad \text{(ii)}\end{aligned}$$

When P is sufficiently above P_{crit}, the sums Σ' are large compared with $|a_0^\circ|^2$ terms on both (i) and (ii), and then it follows that the proportionality

$$\chi''P\propto\langle\delta m_z\rangle$$

is approximately correct.

For the instability at the main resonance, $\omega_k\doteqdot\omega$, the z-directed spin waves are the most important so that $\cosh\psi_k\doteqdot 1$. The proportionality between $\langle\delta m_z\rangle$ and the dissipated power is then exact, provided $\eta_k=\eta_0$. BLOEMBERGEN and WANG[8] give curves of $\langle\delta m_z\rangle$ versus $\chi''P$ at the main resonance, which seem to be quite good straight lines, without any evident break at P_{crit}. This would lead to the conclusion that, indeed $\eta_k=\eta_0$. However, the data on which these straight lines are based are too sparse to really warrant this conclusion.

REFERENCES

1. LANDAU L. and LIFSHITZ E. *Physik. Zeitsch. d. Sowjetunion* **8,** 153 (1935).
2. KITTEL C. *Phys. Rev.* **71,** 270 (1947).
3. VAN VLECK J. H. *Phys. Rev.* **78,** 266 (1950.
4. CLOGSTON A. M. *Bell System Technical Journal* **34,** 739 (1955).
5. AMENT W. S. and RADO G. T. *Phys. Rev.* **97,** 1558 (1955).
6. CLOGSTON A. M., SUHL H., WALKER L. R., and ANDERSON P. W. *Phys. Rev.* **101,** 903 (1956), and paper submitted for publication to the *J. Phys. Chem. Solids.*

VOL. 1 1956/57

7. DAMON R. W. *Rev. Mod. Phys.* **25,** 239 (1953).
8. BLOEMBERGEN N. and WANG S. *Phys. Rev.* **93,** 72 (1954).
9. SUHL H. To be published in *I.R.E. Proceedings* in Oct. 1956. See also ANDERSON P. W. and SUHL H. *Phys. Rev.* **100,** 1788 (1955), and SUHL H. *Phys. Rev.* **101,** 1937 (1956).
10. KASUYA T. *Progress of Theoretical Physics* **12,** 802 (1954).
11. KEFFER F. *Phys. Rev.* **78,** 686 (1952).
12. HERRING C. and KITTEL C. *Phys. Rev.* **81,** 869 (1951).
13. WALKER L. R. *Bulletin of the American Physical Society Pittsburgh Meeting, March* 1956, and a forthcoming publication in *Phys. Rev.*
14. WANG M. C. and UHLENBECK G. E. *Rev. Mod. Phys.* **17,** 323 (1945).
15. DYSON F. J. *Phys. Rev.* **102,** 1217 (1956).
16. HOLSTEIN T. and PRIMAKOFF H. *Phys. Rev.* **58,** 1098 (1940).

J. Phys. Chem. Solids. Pergamon Press 1957. Vol. 1. pp. 228–233.

DETERMINATION OF THE MAGNETIC ANISOTROPY CONSTANT K_2 OF CUBIC FERROMAGNETIC SUBSTANCES

HIROSHI SATO* AND B. S. CHANDRASEKHAR

Westinghouse Research Laboratories, Pittsburgh, Pennsylvania

(*Received* 20 *June* 1956)

VOL. 1 1956/57

Abstract—It is believed, for reasons mentioned in the paper, that the values of the magnetic anisotropy constant of cubic crystals, K_2, hitherto given in the literature are not very reliable. Since K_2 values are often required for the theoretical treatment of the problem, a new, and direct, determination of K_2 for iron and nickel has been made, by means of magnetic torque measurements in the (111) plane. This method minimizes the contribution from K_1 to the torque. Then the pure K_2 contribution is singled out from the measured torque curve with the aid of symmetry considerations. The values obtained for iron and nickel are at room temperature $7{\cdot}14 \times 10^4$ erg/cm^3 and $2{\cdot}34 \times 10^4$ erg/cm^3, and, at liquid nitrogen temperature, $9{\cdot}04 \times 10^4$ erg/cm^3 and $3{\cdot}17 \times 10^5$ erg/cm^3 respectively. The results show, in the case of iron, that the value of K_2 which had been given previously with the same order of magnitude as K_1, is actually an order of magnitude smaller. In the case of nickel, the value is comparable to K_1 and its temperature dependence is even larger than that of K_1.

1. INTRODUCTION

THE determination of the values and temperature coefficients of ferromagnetic anisotropy constants is of importance not only from a practical point of view, but also for theoretical purposes. The theory of ferromagnetic anisotropy is not in a wholly satisfactory state, and an accurate knowledge of the constants is necessary for comparing with theoretical predictions.

The free energy F_k associated with crystalline anisotropy for a cubic crystal is usually written in ascending powers of α's

$$F_k = K_0 + K_1(\alpha_1^2\alpha_2^2 + \alpha_2^2\alpha_3^2 + \alpha_3^2\alpha_1^2) + K_2\alpha_1^2\alpha_2^2\alpha_3^2 + \ldots \quad (1)$$

The α's are the direction cosines of the magnetization vector referred to the cubic axes of the crystal. The K's are the ferromagnetic anisotropy constants. Since the present interest is only in the dependence of the free energy on the direction of magnetization, K_0 will not be considered further.

The magnetization curve in its dependence on direction is dominated by the term in K_1, since the higher terms are more isotropic, and the K's appearing in these higher terms do not affect the magnetization to the same extent. For this reason, the determination of K_1 for different ferromagnetic substances as a function of temperature has received considerable attention, and may be taken to be known with reasonable accuracy. The situation regarding K_2, however, is rather different. There is wide disagreement in the published values, and sometimes even the sign, of K_2 in the literature.[1] This is unfortunate, as a knowledge of K_2, particularly at low temperatures, seems to be necessary to fix the temperature dependence of K_1.

While there is no satisfactory theory of the anisotropy of ferromagnetic metals from a fundamental standpoint, ZENER[2] has been able to correlate formally the temperature dependence of the anisotropy constants with that of the spontaneous magnetization. He uses a simplified model for the mechanism of the temperature dependence of the spontaneous magnetization, and finds that this is related to the anisotropy as follows:

* Now at Scientific Laboratory, Ford Motor Co., Dearborn, Michigan.

$$K_1(T)/K_1(0) \simeq \{J_s(T)/J_s(0)\}^{10}$$

and (2)

$$K_2(T)/K_2(0) \simeq \{J_s(T)/J_s(0)\}^{21}.$$

$K_1(0)$, etc. are here the values of the K's, and $J_s(0)$ is the value of the spontaneous magnetization J_s at the absolute zero of temperature. The above relation for K_1 agrees very well with the experimental data on iron. But in the case of nickel, the observed dependence of K_1 on temperature T is much steeper and the exponent in the relation for $K_1(T)/K_1(0)$ appears to be nearly 20. It is therefore essential in the study of ferromagnetic anisotropy to know what factors affect the temperature dependence of K_1 differently in iron than in nickel. ZENER's considerations are quite formal except for an assumption about the behavior of the magnetization vector with temperature change and, therefore, his conclusions on the temperature dependence of K_1, K_2, etc. should hold at least qualitatively as long as the mechanism of the temperature dependence of the spontaneous magnetization is not much different from that assumed by him. The different effects for iron and nickel might suggest a different mechanism in the temperature dependence of the spontaneous magnetization in iron and nickel respectively. Recently, his assumption was critically examined from a somewhat broader point of view by KEFFER.[3]

On the other hand, a closer examination of ZENER's equations reveals that $K_1(T)$ is given more exactly by the formula

$$K_1(T) = \{K_1(0)+(1/11)K_2(0)\}\{J_s(T)/J_s(0)\}^{10} - -(1/11)K_2(0)\{J_s(T)/J_s(0)\}^{21}+\ldots \quad (3)$$

It is evident from this equation that the value of $K_2(0)$ influences the temperature dependence of K_1, and in particular a sufficiently large value of K_2 with different sign from that of K_1 could cause K_1 to increase with decreasing temperature more rapidly than $\{J_s(T)/J_s(0)\}^{10}$. It is therefore apparent that a knowledge of K_2 is necessary. It was felt, however, for reasons given later, that the values of K_2 given in the literature were not very reliable, and a redetermination of K_2 seemed necessary. Hence the present experiments were carried out with the aim of making as direct a determination of K_2 as possible.

2. PRINCIPLE OF THE MEASUREMENTS

The accurate measurement of K_2 is not easy, as is evidenced by the diversity of values quoted in the literature. The following are the main methods used in the study of ferromagnetic anisotropy:

(i) Measuring the difference in the work of magnetization in different crystallographic directions.

(ii) Comparison of calculated and measured magnetization curves in the principal crystallographic directions.

(iii) measurement of the normal component of the magnetization at high fields.

(iv) From the angular dependence of the internal field coming from the crystalline anisotropy in a ferromagnetic resonance experiment.

Of these, the fourth is a rather indirect method, and the measurement is limited to the surface of the sample, which is likely to be more disturbed than the inside in the course of preparation of the sample. The first two methods, and especially the second, are not very suitable for evaluating K_2, as the terms involving K_2 in the calculation are very small. The accurate determination of the magnetization curve with an exact orientation is a difficult problem, and small errors which are unimportant for K_1 have an excessive effect on the calculated K_2. This leaves only the third method, which has the advantage of utilizing torque measurements which can be made more easily and accurately than magnetization measurements. In addition, the torque, which is proportional to $\mathbf{J_n}\times\mathbf{H}$, where $\mathbf{J_n}$ is the normal component of magnetization in a field $\mathbf{H}$, is expected to be independent of field strength for high enough fields.

These torque measurements have hitherto been done mostly in the (110) plane, using a single-crystal specimen in the form of a flat circular disc. This plane is chosen because it includes the three principal directions of a cubic crystal, and the expression for the torque on the sample in a magnetic field includes terms in both K_1 and K_2. By fitting the experimental torque curve to a theoretical one, one can in principle calculate both K_1 and K_2. But, as has already been pointed out, experimental errors have a much more drastic effect on the calculated K_2 than on K_1, because the coefficient of K_2 is usually one order of magnitude

smaller than that of K_1. The following are some of the errors:

(i) Inaccuracy in the determination (by X-rays, for example) of sample orientation with respect to crystal axes and in the processing of the crystal in this orientation.
(ii) Deviation of sample shape from that ideally assumed.
(iii) Inhomogeneity of magnetization in the sample, which would cause a dependence of the torque on the field.
(iv) Incompleteness of the single crystal used due to e.g., small included grains, included small-angle grain boundary or an additional anisotropy arising from the internal stress. These are almost unavoidable in the case of metallic substances.
(v) Errors in setting the sample in the field.
(vi) Errors coming from the change of the position of the crystal when the field is on, e.g., due to attraction by a pole piece.

If one assumes that these various sources of error contribute an uncertainty of 10 per cent to the value of K_1 (in fact the scatter in the values of K_1 hitherto reported indicates a larger uncertainty), and if $K_2 \approx K_1$, then the error in K_2 would be of the order of 100 per cent. It is thus evident that, although the method can in principle give K_2, in fact the results would be unreliable.

Fortunately, there is one experimental configuration which is very favorable for the evaluation of K_2. As one can see from equation (1), K_1 does not enter into the equation for the torque curve in the (111) plane, which can be written in the form

$$-\partial F_K/\partial\theta = -K_2 \cdot (\sin 6\theta)/18 \qquad (4)$$

neglecting higher terms, where θ is measured from the $\langle 110\rangle$ direction. Thus, if the torque can be accurately measured in the (111) plane, the torque curve should have a (6θ) dependence and from its amplitude one can easily obtain the value of K_2 according to equation (4).

Even in this plane, the result cannot be free from the errors pointed out above which may cause the K_1 contribution, the coefficient of which is one order of magnitude larger than that of K_2, to mask the K_2 contribution. However, one has reason to expect that the contributions from the main sources of error will in practice be only up to terms in $\sin 4\theta$ and $\cos 4\theta$. The most probable systematic deviations from the ideally assumed symmetry of the sample, for example, a uniform variation in thickness from one edge to the other, will contribute terms in $\sin 2\theta$ and $\cos 2\theta$, while errors, due to the magnetic field not being exactly parallel to the (111) plane, give terms up to $\sin 4\theta$ and $\cos 4\theta$.[4]

In the actual case, however, it should be noticed that the magnetization generally does not lie in the plane of the disc ((111) plane) as assumed above, but oscillates about the plane with a six-fold symmetry (being zero in $\langle 110\rangle$ directions and maximum in $\langle 112\rangle$ directions) during the one cycle of the experiment in the case of both iron and nickel, due to the geometrical relation between the magnetic field and the direction of easy magnetization, $\langle 100\rangle$ or $\langle 111\rangle$. The amplitude of this oscillation decreases as the magnetic field increases. Because of this fact, the specimen receives another torque of sixfold symmetry originating from the K_1 term of the anisotropy even if the plane of the disc exactly coincides with a (111) plane. The magnitude of this torque decreases as the magnetic field increases becoming zero in both the $\langle 110\rangle$ and the $\langle 112\rangle$ directions. However, at the field of 16,000 oe used, the deviation of the magnetization from the plane of the disc is easily estimated to be, at most, of the order of ten minutes, and the magnitude of the extra torque of six-fold symmetry due to the K_1 term would then be less than $K_1/1000$ and practically negligible. This is borne out also by the observation that the torque was independent of the field in our measurements.

The other sources of errors should be quite irregular. The 6θ component of the torque curve is, therefore, entirely from the contribution of K_2, because the slight deviation of the plane from the exact orientations would not affect the K_2 contribution much. These considerations enable us to obtain K_2 from the torque curves from which, otherwise, it would have been rather difficult to get any reliable information about its value.

3. PREPARATION OF SAMPLES AND THE MEASURING APPARATUS

Iron and nickel single-crystal specimens were used. The iron single crystals were prepared by the strain-annealing method from Armco iron. By the combination

VOL. 1 1956/57

of appropriate preparatory heat and mechanical treatment followed by strain-annealing, it was very easy to grow rod single crystals of iron more than 1 in. in diameter and 2 in. in length, but it was almost impossible to avoid having very small grains dispersed in them, especially on their surface.[5] Nickel single crystals were prepared from vacuum melted electrolytic nickel by the Tamman-Bridgman method. In the case of iron, good parts which did not include appreciable grains were selected for preparing the specimens.

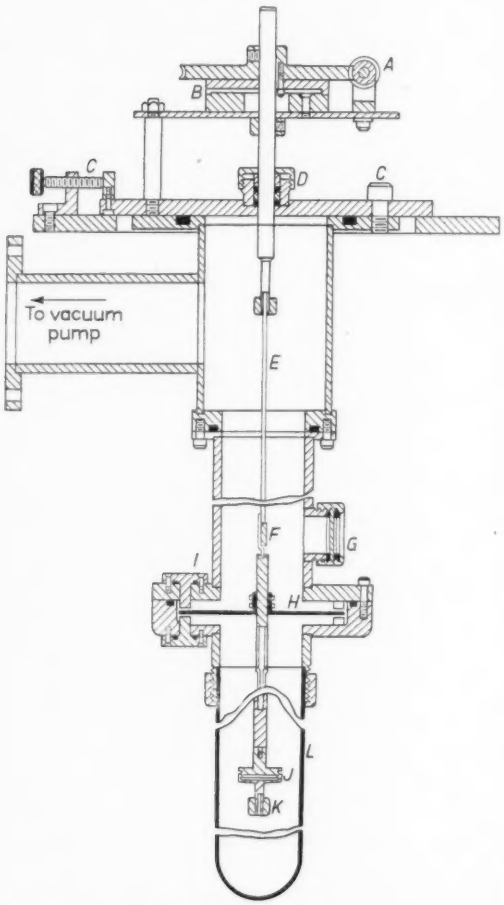


FIG. 1. General Assembly of the Measuring Apparatus.

A: Worm gear, reduction ratio 30 : 1
B: Scale and vernier
C: Centering and leveling screws
D: Double "O" ring seal for the stem
E: Phosphor-bronze suspension wire
F: Mirror
G: Glass window
H: Damping fin
I: Pole pieces of an electromagnet for damping device
J: Case for holding sample
K: Pin vice for weights
L: Quartz tubing.

The crystals were first cut on a milling cutter after a rough orientation determination by the back reflection Laue method with an error of the order of 1° and then a precision orientation was done using an X-ray spectrometer with a jig similar to the one used by WALKER and others.[6] Particulars of this apparatus will be published elsewhere.[7] The final specimens were circular discs of diameter approximately 1/2 in. and thickness ranging from 0·02 in. to 0·1 in.

The apparatus for the measurement was essentially a torsion balance and the general assembly is shown in Fig. 1. The inside of the apparatus can be evacuated and is designed to measure the properties down to liquid-helium temperatures. The suspension wire can be rotated through a double "O" ring seal, by a reduction gear motor and worm gear. The reduction ratio of the worm gear is 30 and the r.p.m. of the reduced gear motor is 2, giving 24°/min for the speed at which the torsion head can be turned. The head is graduated and can be read through a telescope and a vernier to within $\pm 0{\cdot}1°$. The torque is determined, using a null method, from the angle of rotation read at the head. The suspension wire is well separated from the specimen and is always kept at room temperature. Because of this, the torque constant of the suspension wire is independent of the changing temperature of the specimen. An Arthur D. Little magnet is used, which provides enough field strength even in the measurement at liquid helium temperature, where because of the increase of K's on the one hand and the space necessary for double Dewar vessels on the other, a high homogeneous magnetic field at large pole piece distances is necessary.

Because we use a type of torsion balance where the sensitivity is proportional to $1/r^4$, r being the diameter of the suspension wire, we can get almost any desired sensitivity. Phosphor bronze wires of 32 mil to 8 mil diameter were used for the suspension enabling the sensitivity to be changed by a factor of up to 250, as needed. Damping of the balance is automatically obtained, because weights of copper or brass are used to prevent the specimen from sticking to the pole pieces; and the oscillation of the weights in a strong magnetic field provides the necessary damping. In addition, an electromagnetic damping system is provided just below the suspension.

4. RESULTS OF THE MEASUREMENT

Results of the measurements at room and liquid nitrogen temperatures are discussed here. The magnetic field used was about 16,000 oe. In this field it can be safely assumed in the analysis that the directions of the magnetization and the field are practically the same. The measured torque was indeed found to be field-independent at this field strength. As an example, a measured torque curve in a (111) plane of a nickel single crystal at room temperature is shown in Fig. 2. In this case, the K_2 contribution relative to that of K_1 is rather

large and one can clearly observe the component with a 60° period that was expected. In the case of iron, on the other hand, such a 60° modulation is not very apparent. This means that the value of K_2 in the case of iron is rather small compared to

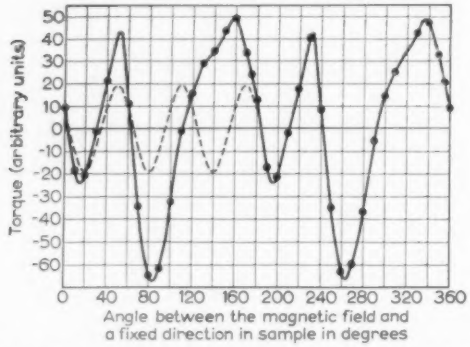


FIG. 2. Torque curve of a single crystal disc of nickel in the (111) plane, taken at room temperature (26°C) with an applied field of 16,000 oersteds. The broken line is the calculated K_2 contribution. One unit here corresponds to a torque of 4·442 dyne-cm or 68·1 dyne-cm/cm³ of sample in this case.

the value of K_1. Even in the case of nickel, though the deviation of the specimen from an exact (111) plane is very small, the K_1 and other unwanted contributions coming from several sources are remarkable. For the evaluation of K_2, therefore, the K_2 contribution should be separated clearly from the experimental curve. This can be done because only the K_2 contribution has the period of 6θ as explained before. It is necessary therefore to analyse the torque curve to extract the 6θ component. For this purpose, however, one must know the orientation of the crystal with respect to the magnetic field, in order to fix the zero point of the (6θ) component. The determination of the angle between a definite direction in the crystal and the field direction can be done, but only with considerable error. This difficulty can be avoided in the following way. According to equation (4) the K_2 contribution is expressed by a single 6θ sinusoidal curve and utilizing this fact one can get the zero point as follows:

If one analyses harmonically the experimental curve, taking different positions as the origin, one can expect a dependence of the coefficient of sin 6θ or cos 6θ on the position of the origin. If one plots these values against the positions of the origin, one gets a curve, an example of which is shown in Fig. 3 in the case of nickel at room temperature shown in Fig. 2. The point where the coefficient becomes a maximum is the point where the sin 6θ or cos 6θ curve starts and the maximum value gives the amplitude of the curve. This method also

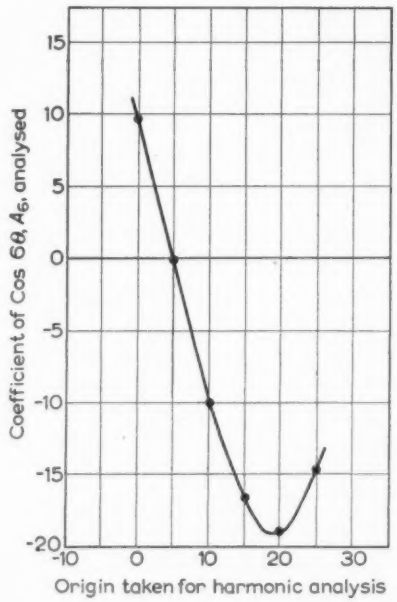


FIG. 3. A dependence of the coefficient of cos 6θ on the position of the origin taken for the harmonic analysis of a torque curve. The torque curve is that for nickel at room temperature shown in Fig. 2. The units for the ordinates and the abscissae are the same as in Fig. 2.

averages out some of the errors that are inherent in computing the harmonic coefficients with a finite number of terms, by taking a different combination of points on the curve each time. From Fig. 3, one can clearly see that the coefficient increases from zero to a maximum in 15°. Thus one can determine the origin as well as the orientation of the crystal with respect to the magnetic field, if an approximate orientation is known beforehand, and the amplitude of the 6θ curve. This origin gives the sign, and the amplitude gives the value of K_2 according to equation (4). This procedure enables K_2 to be determined with little ambiguity—something that hitherto had been considered very difficult.

The values for iron and nickel are given in Table 1. It is immediately apparent that these values are very much different from what had been obtained by earlier workers. The value for iron is almost an order of magnitude smaller than some of the earlier values. The K_2 for nickel is

VOL. 1 1956/57

Table 1. *The values of* K_2 *at room and liquid-nitrogen temperature*

	Room Temperature (26°C)		Liquid-Nitrogen Temperature (−196°C)	
	K_1(erg/cm³)	K_2(erg/cm³)	K_1(erg/cm³)	K_2(erg/cm³)
Fe	$4{\cdot}5\times10^5$	$7{\cdot}14\times10^4$	$5{\cdot}7\times10^5$	$9{\cdot}04\times10^4$
Ni	$-4{\cdot}5\times10^4$	$2{\cdot}34\times10^4$	$-5{\cdot}7\times10^5$	$3{\cdot}17\times10^5$

The values of K_1 were obtained from BOZORTH's book.[1]

relatively large; this had been, however, expected from the behavior of the magnetization curve. An interesting point to be noted here is that the temperature dependence of K_2 in nickel seems to be even larger than that of K_1 and its value is quite large at liquid nitrogen temperature. These results are not inconsistent with ZENER's theory. The tenth power law holds very well experimentally in the case of iron. This is expected on his theory only when $K_2(0)$ is rather small compared to $K_1(0)$ and the experimental results show that this is the case.

In the case of nickel, the signs and magnitudes of K_1 and K_2 are such as to predict a greater temperature dependence of K_1, although the magnitude of the theoretical effect predicted by equation (3) is inadequate to explain the observed dependence of K_1 on temperature.

Experiments to give data covering a wider temperature range are under way.

5. CONCLUDING REMARKS

The determination of the second coefficient of ferromagnetic anisotropy, whose values previously determined had shown much scatter, was reattempted by a torque measurement in the (111) plane. The fact that the unwanted contribution to the torque curve is only in terms up to sin 4θ and cos 4θ makes possible a reasonably accurate analysis of K_2. The results show, in the case of iron, that the value of K_2, which had been given previously with the same order of magnitude as K_1, is actually an order of magnitude smaller. In the case of nickel, the value is comparable to K_1 and its temperature dependence is even larger than that of K_1. Both these facts agree qualitatively with the predictions of ZENER's treatment of magnetic anisotropy.

Acknowledgements—The authors wish to express their thanks to Dr. C. ZENER for his interest and encouragement. Thanks are also due to Mr. J. A. OSBORN, Dr. A. TAYLOR, and Dr. B. A. CALHOUN for their assistance in several stages of the experiment.

REFERENCES

1. BECKER R. and DOERING W. *Ferromagnetismus* Julius Springer, Berlin, Germany, 112 (1938).
   BOZORTH R. M. *Ferromagnetism*, D. Van Nostrand Co., New York, 567 (1951).
2. ZENER C. *Phys. Rev.* **96**, 1335 (1954).
3. KEFFER F. *Phys. Rev.* **100**, 1692 (1955).
4. TARASOV L. P. and BITTER F. *Phys. Rev.* **52**, 353 (1937).
5. SATO H. Research Report, 8-0107-R3, Westinghouse Res. Labs. (1956).
6. WALKER J. T., WILLIAMS H. J., and BOZORTH R. M. *Rev. Sci. Instrum.* **20**, 947 (1949).
7. SATO H., GETSKO J., and HICKMAN W. E. *Rev. Sci. Instrum.*, to be published.

J. Phys. Chem. Solids. Pergamon Press 1957. Vol. 1. pp. 234–248.

ANALYSIS OF PHOTOCONDUCTIVITY APPLIED TO CADMIUM-SULFIDE-TYPE PHOTOCONDUCTORS

RICHARD H. BUBE

RCA Laboratories, Radio Corporation of America, Princeton, New Jersey

(*Received* 27 *July* 1956)

Abstract—The model of a photoconductor including two different types of recombination centers, filled in the dark, and with trapping centers, empty in the dark, is able to adequately describe such photoconductivity phenomena in CdS-type photoconductors as supralinearity, temperature dependence, infra-red quenching, and variations in speed of response. Thus the conceptual description of ROSE, which has been previously shown to have semiquantitative validity, is put on a more detailed quantitative basis. The application of the theoretical calculations to experimental data for CdSe gives self-consistent values for the ratio (8×10^5) of the capture cross-section of the "sensitizing" centers for holes to that for electrons, and for the energy difference (0·64 eV) between the level corresponding to this type of center and the top of the valence band. On the basis of the model, the relatively fast response of CdSe is associated with the specific location of these levels, which location carries with it the occurrence of supralinearity and large temperature dependence of photosensitivity.

VOL. 1 1956/5

INTRODUCTION

RECENT investigations of a number of different photoconductivity phenomena (supralinearity, temperature dependence, infra-red quenching, speed of response) in CdS-type photoconductors[1–4] have shown the semiquantitative validity of the conceptual model of a photoconductor proposed by ROSE.[5–8] The essence of this model is the existence in the photoconductor of two different types of recombination centers, as characterized by their capture cross-sections. Illustrative calculations on similar models have been reported by DUBOC,[9] DIRKSEN and MEMELINK[9a] and SCHÖN.[10] Recently almost identical phenomena have been reported for germanium with divalent impurities.[11, 12]

It is the purpose of this paper to explore in somewhat greater detail the mathematical analysis and the physical interpretation of a representative type of such a model. The insight obtained from a consideration of the model may then be used to analyse data of photocurrent as a function of light intensity, measured at various temperatures, to obtain numerical values for the energy separation between the levels of the low-lying "sensitizing" centers and the top of the valence band, and also the ratio of the capture cross-section of these centers for holes to that for electrons.

Additional investigation of photosensitivity and speed of response as a function of temperature for CdS and CdSe indicates that on the basis of the model, the relatively fast response of CdSe is associated with the specific location of these "sensitizing-center" levels, which location (for

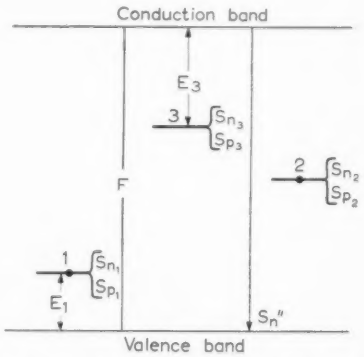


FIG. 1. Energy level scheme for model of a photoconductor.

normal operation based at room temperature) has of necessity connected with it both marked temperature dependence of photosensitivity and speed of response, and supralinear photocurrent vs. light-intensity variation.

THEORY

The model to be considered is shown in Fig. 1. Three sets of levels are assumed to exist in the forbidden gap; levels of types 1 and 2 are assumed filled in the dark, and levels of type 3 are assumed empty in the dark. The symbols to be used have the following definitions.

N_i — density of states i
n_i — density of filled states i
S_{n_i} — capture cross-section of states i for electrons
S_{p_i} — capture cross-section of states i for holes
n — density of free electrons
p — density of free holes
S_n'' — recombination cross-section between free electrons and free holes
F — density of electron-hole pairs created per second by optical excitation across the band gap
β_{n_i} — product of S_{n_i} and the thermal velocity for an electron
β_{p_i} — product of S_{p_i} and the thermal velocity for a hole
β_n — product of S_n'' and the thermal velocity
E_1 — energy difference between levels 1 and top of valence band
E_3 — energy difference between bottom of conduction band and levels 3
P_1 — probability per unit time for thermal ejection of a hole in levels 1 into the valence band

$$P_1 = N_v \beta_{p_1} \exp(-E_1/kT)$$

where N_v is the effective density of states in the uppermost kT-wide part of the valence band given by:

$$N_v = 2\left(\frac{2\pi m_p kT}{h^2}\right)^{3/2}$$

where m_p is the effective mass for holes.

P_3 — probability per unit time for thermal ejection of an electron in levels 3 into the conduction band

$$P_3 = N_c \beta_{n_3} \exp(-E_3/kT)$$

where N_c is the effective density of states in the lowest kT-wide part of the conduction band defined analagously to N_v.

The equation for charge neutrality is:

$$n+n_3 = p+\sum_{i=1}^{2}(N_i-n_i). \tag{1}$$

The following five equations express the rate of change of the various densities involved:

$$dn/dt = 0 = F-n\sum_{i=1}^{3}\beta_{n_i}(N_i-n_i)+n_3P_3-\beta_n''np \tag{2}$$

$$dp/dt = 0 = F-p\sum_{i=1}^{3}\beta_{p_i}n_i+(N_1-n_1)P_1-\beta_n''np \tag{3}$$

$$dn_1/dt = 0 = (N_1-n_1)P_1-\beta_{p_1}n_1p+\beta_{n_1}n(N_1-n_1) \tag{4}$$

$$dn_2/dt = 0 = \beta_{n_2}n(N_2-n_2)-\beta_{p_2}n_2p \tag{5}$$

$$dn_3/dt = 0 = \beta_{n_3}n(N_3-n_3)-\beta_{p_3}n_3p-n_3p_3. \tag{6}$$

The rates have been set equal to zero in the above equations to give the conditions for the steady state. (Only four of the equations (2) through (6) are independent.)

Solutions for the n_i in terms of n and p can be obtained from equations (4) through (6):

$$n_1 = N_1\{1+\beta_{p_1}p/(\beta_{n_1}n+P_1)\}^{-1} \tag{7}$$

$$n_2 = N_2(1+\beta_{p_2}p/\beta_{n_2}n)^{-1} \tag{8}$$

$$n_3 = N_3\{1+(\beta_{p_3}p+P_3)/\beta_{n_3}n\}^{-1}. \tag{9}$$

Using equations (7) through (9) together with equation (1), p can be expressed in terms of n.

$$p = n-N_1-N_2+N_1\{1+\beta_{p_1}p/(\beta_{n_1}n+P_1)\}^{-1}+ +N_2(1+\beta_{p_2}p/\beta n_2n)^{-1}+ +N_3\{1+(\beta_{p_3}p+P_3)/\beta_{n_3}n\}^{-1}. \tag{10}$$

Equations (7) through (9) can be substituted into equation (3) to give F in terms of n and p:

$$F = p[\beta_{p_1}N_1\{1+(\beta_{p_1}p+P_1)/\beta_{n_1}n\}^{-1}+ +\beta_{p_2}N_2(1+\beta_{p_2}p/\beta_{n_2}n)^{-1}+ +\beta_{p_3}N_3\{1+(\beta_{p_3}p+P_3)/\beta_{n_3}n\}^{-1}+\beta_n''n]. \tag{11}$$

Equation (10) can then be used to eliminate p from equation (11) and thus express F in terms of n. The actual substitution was not carried out in detail because of the obvious complexity which the results would have, but calculations were made to solve the two equations arithmetically, thus preserving the possibility of understanding the variations physically.

In the consideration of the dynamics of the photoconductivity process, it is frequently convenient to describe phenomena in terms of the quasi-Fermi levels and the demarcation lines corresponding to the various centers. The location of the quasi-Fermi levels and demarcation lines for the conditions of low-light intensity and high-light intensity are indicated schematically in Fig. 2.

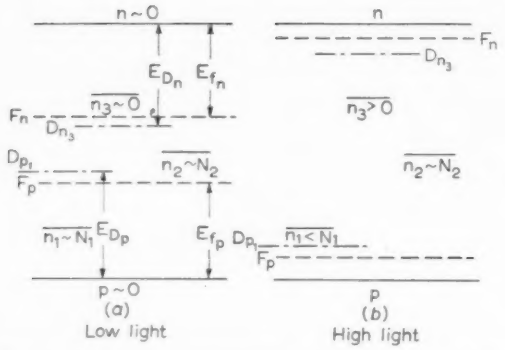


FIG. 2. Schematic representation of the increase in sensitivity of a photoconductor between low light and high light excitation intensity.

There are five equations relating the quasi-Fermi levels and the demarcation lines which it is helpful to keep in mind:

$$E_{Fn} = E_{D n_2} + kT \ln(S_{p_2} p / S_{n_2} n) \tag{12}$$

$$E_{Fp} = E_{D p_1} - kT \ln(S_{p_1} p / S_{n_1} n) \tag{13}$$

$$E_{D p_1} = E_{Fn} + kT \ln(S_{p_1} / S_{n_1}) \tag{14}$$

$$E_{D n_2} = E_{Fp} - kT \ln(S_{p_2} / S_{n_2}) \tag{15}$$

$$E_{Fp} = E_{Fn} - kT \ln(p/n). \tag{16}$$

These equations as written above contain the assumption that $N_c = N_v$. They are also based on the definition of the quasi-Fermi levels:

$$E_{Fn} = kT \ln(N_c/n) \tag{17}$$

$$E_{Fp} = kT \ln(N_v/p). \tag{18}$$

The meaning of the demarcation line is the following: When an electron (hole) is located at the electron (hole) demarcation line, it has equal probability of recombining with a free hole (electron) and of being thermally ejected to the conduction (valence) band.

The relation between the quantities of equations (12) through (18) and those of equations (10) and (11) can be illustrated by four examples: (1) When $\beta_{n_2} n$ is equal to P_3, under conditions such that the centers 3 are still in thermal equilibrium with the conduction band as far as capture and ejection of electrons are concerned, the electron quasi-Fermi level, F_n, is located at the levels 3; (2) when $\beta_{p_2} p$ is equal to P_3, the electron demarcation line, D_{n_2}, for the centers 3 is at the levels 3; (3) when $\beta_{p_1} p$ is equal to P_1, under conditions such that the centers 1 are still in thermal equilibrium with the valence band, the hole quasi-Fermi level, F_p, is located at the levels 1; (4) when $\beta_{n_1} n$ is equal to P_1, the hole demarcation line, D_{p_1}, for the centers 1 is at the levels 1. It is clear that at least some of the changes in the variation of n with F will occur at concentrations of n and p which are described in these examples.

In order to make the arithmetical calculations on the basis of equations (10) and (11) somewhat less complicated, certain simplifying assumptions were made, as well as the assumption of reasonable numerical values for the various parameters.

$$\left.\begin{aligned}
&N_1 = N_2 = 10^{15}/\text{cm}^3 \\
&S = S_{p_1} = S_{p_2} = S_{p_3} \\
&\quad = S_{n_2} = S_{n_3} = 10^{-15}\text{cm}^2; \; Sv = \beta \\
&S' = S_{n_1} = 10^{-19}\text{cm}^2 \qquad ; S'v = \beta' \\
&S_n'' = 10^{-20}\text{cm}^2 \qquad ; S_n'' v = \beta'' \\
&P_1 = 10^{-1}\ \text{sec}^{-1} \text{ at } 300^\circ\text{K} \\
&\quad (\text{corresponds to } E_1 = 0{\cdot}7\ \text{eV}) \\
&P_3 = 10^{-4}\ \text{sec}^{-1} \text{ at } 300^\circ\text{K} \\
&\quad (\text{corresponds to } E_3 = 0{\cdot}9\ \text{eV}).
\end{aligned}\right\} \tag{19}$$

Calculations were carried out for $N_3 = 10^{17}$, 10^{15}, 10^{13}, and 10^7 cm^{-3}. The assumptions used imply principally that centers 2 and 3 are of the same type with about equal cross-sections for capture of either electrons or holes; they could, therefore, be replaced in the model by a single set of centers which are only partially filled in the dark.* Centers 1, on the other hand, have a cross-section for holes the same as that of centers 2 and 3, but have a considerably smaller cross-section for electrons.

The calculation for the case of $N_1 = N_2 = N_3 = 10^{15}\text{cm}^{-3}$ leads to the variation of n and p as a function of F shown in Fig. 3. The corresponding variation of n_1, n_2, and n_3 with F are shown in Fig. 4. Most of the essential features of the dependence of photocurrent on light intensity are found in Fig. 3. At low values of F, where the quasi-Fermi level for electrons and the electron demarcation line for centers 3 lie below the levels 3, (and of course the quasi-Fermi level for holes and the

* An equivalent way of considering these states would be to assume a continuous distribution of levels corresponding to centers of type 2 with the dark electron Fermi level located within this continuous distribution. The only result of the present discussion which would be appreciably altered by such an assumption would be the low-light dependence of n on F.

VOL. 1 1956/5

hole demarcation line for centers 1 lie above the levels 1), the density of free electrons varies as the square root of the light intensity. In this range of F, the density of trapped electrons in centers 3 is essentially the same as the density of empty recombination centers (centers 2), and the density of free electrons is proportional to the density of trapped electrons.

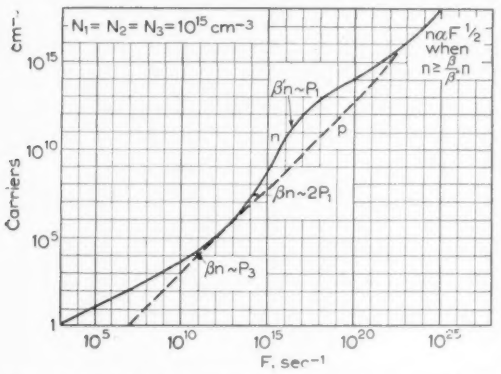


FIG. 3. Calculated density of charge carriers as a function of light intensity for $N_1 = N_2 = N_3 = 10^{15}/\text{cm}^3$.

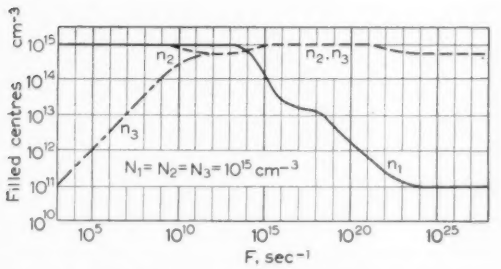


FIG. 4. Calculated density of filled centers as a function of light intensity for $N_1 = N_2 = N_3 = 10^{15}/\text{cm}^3$.

When βn (and for slightly higher F, βp) exceeds P_3, the centers 3 become about half-occupied, the electrons initially in the centers 2 being equally divided between centers 2 and centers 3. In this range of F, where the density of empty centers 1 is still negligible compared to the density of empty centers 2 and 3, the density of empty recombination centers is essentially constant, and the density of free electrons is proportional to the light intensity.

As the light intensity is increased further, the point will be reached where the photocurrent will begin to rise supralinearly with light intensity if the conditions are suitable. The conditions for the onset of supralinearity can be obtained in the following way. Consideration of equation (11) indicates that except for slight changes of the order of two, F will vary linearly with p, and that, therefore, the variation of n with F is determined approximately by the variation of p with n. Examination of equation (10) under the conditions which will apply for the onset of supralinearity, i.e., $\beta p \gg P_3$, $\beta n \gg P_3$, and $\beta' n \ll P_1$, gives for $p \ll n$, and $\beta p \ll P_1$†

$$p = nP_1(n+N_3)\{n\beta N_1+(n+N_2+N_3)P_1\}^{-1}. \quad (20)$$

In this formulation, the simplifying assumptions of equations (19) have been used, but unequal N_1, N_2, and N_3 are allowed. The n vs. F variation will be sublinear if p varies with a power of n greater than unity, linear if p varies linearly with n, and supralinear if p varies with a power of n less than unity.

Under the condition that $n \ll N_1, N_2, N_3$, equation (20) reduces to:

$$p = nN_3P_1\{n\beta N_1+(N_2+N_3)P_1\}^{-1}. \quad (21)$$

If $(N_2+N_3)P_1 \gg n\beta N_1$, p varies linearly with n, and n varies linearly with F. If, however, $(N_2+N_3)P_1 \ll n\beta N_1$, p is constant and n varies supralinearly with F. Thus the condition for the onset of supralinearity is given by:

$$\beta n = (N_2+N_3)P_1/N_1. \quad (22)$$

If n is not $\ll N_1, N_2, N_3$, supralinearity may not set in under certain conditions. As one example of such a case, consider $n < N_2$ but $n > N_3$. Then equation (20) becomes:

$$p = n^2P_1(n\beta N_1+N_2P_1)^{-1}. \quad (23)$$

† In the case where $P_3 \gg \beta n, \beta p$, i.e., the levels 3 lie near the conduction band, supralinearity does not occur. It may be shown that a variation $p \propto n^2$ changes to $p \propto n$ when $\beta n = (N_3/N_1)P_1$.

In the case where $\beta p \gg P_1$, an investigation of the conditions for the possible onset of supralinearity show that the assumption of $\beta p \gg P_1$ holds only for $N_3 > N_1$. Supralinearity does not occur; a variation $p \propto n$ changes to $p \propto n^2$ when $n = (N_3-N_1)$. If $\beta p \ll P_1$ at the onset of supralinearity, it is, of course, possible for βp to become $\gg P_1$ during the supralinear range (see equation 26).

If $N_2P_1 \gg n\beta N_1$, p varies as n^2, and n varies as $F^{\frac{1}{2}}$; if $N_2P_1 \ll n\beta N_1$, p varies as n, and n varies as F. Thus supralinearity does not occur. Supralinearity will not occur in general if $n > N_3$.

It may be shown that the condition for the onset of supralinearity given in equation (22) (for $\beta p \gg P_3$, $\beta n \gg P_3$) corresponds to the situation where the density of empty centers 1 is equal to the density of empty centers 2 and 3: $(N_1-n_1) = (N_2-n_2)+(N_3-n_3)$. The general condition for the onset of supralinearity may be stated as follows: supralinearity sets in when the number of electrons removed from centers 1 becomes comparable to the number of holes in centers 2 and 3 below the corresponding electron demarcation lines, provided that the number of free electrons at that condition is less than the number of electrons transferred from center 1 to centers 2 and 3.

Equations (20) through (23) have been based on the assumption that $p \ll n$. If the case for $p \gg n$ is considered, as will occur if $N_3 \gg N_1$, N_2, it is found that p will vary linearly with n:

$$p = nN_3/(N_1+N_2). \tag{24}$$

Supralinearity does not occur for $p \gg n$, the asymmetry being due to the assumption that $S_{n_1} \ll S_{n_2}$, S_{n_3}; $S_{p_1} = S_{p_2} = S_{p_3}$.

Referring back to the calculation summarized in Figs. 3 and 4, it should be noted that recombination of electrons and holes is always dominated by the centers 2 and 3 because of the small value assumed for S_{n_1}, under the other assumptions used. Maximum sensitivity (n/F) occurs for $n = 10^{12}/\text{cm}^3$, for which the density of empty centers 1 is almost $10^{15}/\text{cm}^3$ and the density of empty centers 2 and 3 is a total of $1{\cdot}4 \times 10^{13}/\text{cm}^3$. But because of the difference in the capture cross-sections for electrons, more than 99 per cent of the recombination traffic still takes place through centers 2 and 3.

When the light intensity, F, has increased to the point where $\beta' n = P_1$, the hole demarcation line of centers 1 is at the levels 1. It would be expected conceptually that this situation would mark the end of the supralinear range, and a consideration of the appropriate calculations confirms this fact. Under assumption that $\beta p \gg P_3$, $\beta n \gg P_3$, $p \ll n$, and $n \ll N_1$, N_2, N_3, it may be shown that if $\beta p \ll (\beta' n+P_1)$, (as will occur for $N_3 \ll N_1$, N_2):

$$p = nN_3(\beta' n+P_1)\{n\beta N_1+(\beta' n+P_1)(N_2+N_3)\}^{-1} \tag{25}$$

or that if $\beta p \gg (\beta' n+P_1)$, (as will occur if $N_3 \approx N_1$, N_2):

$$p = \{4\beta n N_1(N_2+N_3)(\beta' n+P_1)\}^{\frac{1}{2}}/\{2\beta(N_2+N_3)\}. \tag{26}$$

An invariance of p with n (eq. 25), or a variation of p with $n^{\frac{1}{2}}$, (eq. 26), for $\beta' n \ll P_1$, both change to a variation of p linearly with n for $\beta' n \gg P_1$.†

As the light intensity is raised further beyond the end of the supralinear range, a range is first encountered where the density of empty recombination centers is approximately constant, and the density of free electrons varies linearly with light intensity. Then for a range of higher light intensities, the density of empty recombination centers increases (from $7 \times 10^{12}/\text{cm}^3$ to $7 \times 10^{14}/\text{cm}^3$, and hence not discernible in Fig. 4), and the density of free electrons varies sublinearly with light intensity. With continued increase of light intensity, the occupancy of centers 1 drops to that given by the ratio of its capture cross-sections, i.e., to 10^{-4}, when $n = p$. Similarly, the occupancy of centers 2 and 3 becomes stabilized at 0·5, since the cross-sections of these centers for electrons and holes are equal. Until direct recombination between free electrons and free holes becomes important, therefore, the density of recombination centers is constant, and the density of free electrons varies linearly with the light intensity. Finally, when $n > (\beta/\beta'')N$, direct recombination between free electrons and holes becomes the major recombination process, and the density of free electrons varies as the square root of the light intensity.

Fig. 5 illustrates the case discussed in connection with equation (24), with $N_3 \gg N_1$, N_2 and $p \gg n$. Supralinearity is not found.

Fig. 6 illustrates the case where $N_3 \ll N_1$, N_2, but is still large enough so that equation (22) applies and not equation (23).

† It is of interest to note that equation (25) describes the case where the supralinear range occurs for both hole demarcation line for center 3 and hole Fermi level above the level 3. Equation (26), on the other hand, describes the case where the supralinear range occurs for the hole demarcation line for centers 3 above the levels 3, but the hole Fermi level below the levels 3. In the former case, the hole demarcation line passes through the levels 3 before the hole Fermi level; in the latter case the reverse is true.

VOL. 1 1956/57

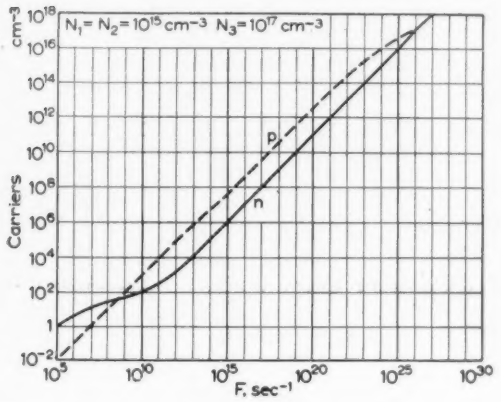


FIG. 5. Calculated density of charge carriers as a function of light intensity for $N_1 = N_2 = 10^{15}/\text{cm}^3$; $N_3 = 10^{17}/\text{cm}^3$.

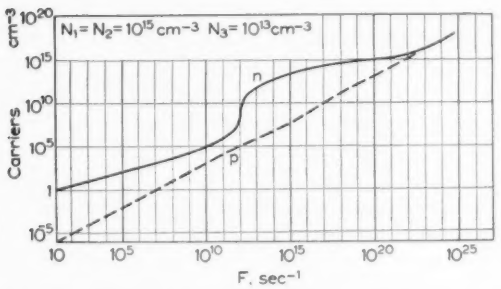


FIG. 6. Calculated density of charge carriers as a function of light intensity for $N_1 = N_2 = 10^{15}/\text{cm}^3$; $N_3 = 10^{13}/\text{cm}^3$.

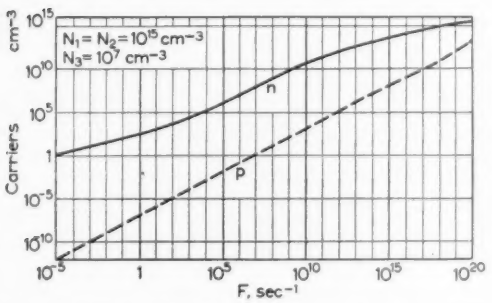


FIG. 7. Calculated density of charge carriers as a function of light intensity for $N_1 = N_2 = 10^{15}/\text{cm}^3$; $N_3 = 10^7/\text{cm}^3$.

Finally Fig. 7 illustrates the case where N_3 is so much smaller than N_1 and N_2 that equation (23) holds and no supralinear region is found.

Two additional facts may be noted: (1) the ratio of n/p for any value of F increases continuously with decreasing N_3 relative to N_1 and N_2; (2) in the case where N_3 is very small and supralinearity is not found, recombination traffic for certain ranges of light intensity can be shared by centers 1 and centers 2 and 3; in the limiting case of $N_3 = 0$, recombination traffic is shared equally between centers 1 and centers 2 for moderately high F.

All of the above calculations have shown the changes in photosensitivity possible at fixed temperature with increasing light intensity. Similar changes in photosensitivity can be caused by varying the temperature, thus changing P_1 and P_3, while keeping the light intensity constant. Fig. 8 illustrates the case for $N_1 = N_2 = N_3 = 10^{15}/\text{cm}^3$ and $F = 10^{16}$ sec^{-1}. The lifetime of a free electron, t^*, can also be calculated (for $\beta p \gg P_3$, $\beta n \gg P_3$):

$$t^* = \{\beta'(N_1 - n_1) + \beta(N_2 - n_2) + \beta(N_3 - n_3)\}^{-1}. \quad (27)$$

The lifetimes calculated in this way for the case of Fig. 8 are shown by the scale on the right of the figure; the same curve is applicable to both the variation of n and t^* with temperature.

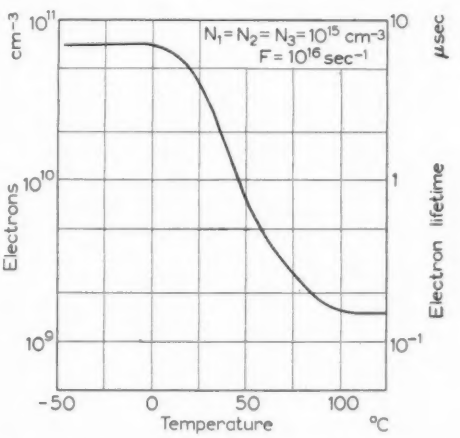


FIG. 8. Density of electrons and electron lifetime as a function of temperature, calculated for $N_1 = N_2 = N_3 = 10^{15}/\text{cm}^3$; $F = 10^{16}$/sec.

The variation of photosensitivity with temperature can be summarized by considering three temperatures: (1) a low temperature for which βn, $\beta p \gg P_3$; βn, $\beta p \gg P_1$; and $\beta' n \gg P_1$; (2) an intermediate temperature for which βn, $\beta p \gg P_3$; βn, $\beta p \ll P_1$; and $\beta' n \ll P_1$; and (3) a high temperature for which βn, $\beta p < P_3$; βn, $\beta p \ll P_1$, and $\beta' n \ll P_1$. For temperatures (1), the crystal is sensitive, with the hole demarcation line for centers 1 lying below the levels 1 and the electron demarcation line for centers 3 lying above levels 3. For temperatures (2), the crystal has passed through a range of temperature quenching and has lost sensitivity; the hole demarcation line for centers 1 lies above the levels 1, but the electron demarcation line for centers 3 still lies above the levels 3. For temperatures (3), centers 3 are being thermally emptied, the electron demarcation line for centers 3 lies below levels 3. Only the range from temperatures (1), through the temperature quenching, to the beginning of temperatures (2) is shown in Fig. 8.

With the assumptions that $p \ll n$, and that $n \ll N$, values of n and p in these ranges can be calculated for the case where $N = N_1 = N_2 = N_3$ as follows:

Low Temperature

$$p = F/(2\beta N) \tag{28a}$$

$$n = (F/N)(2\beta\beta')^{-\frac{1}{2}}, \text{ for } 2\beta nN \ll F \ll 2\beta N^2 \tag{28b}$$

Intermediate Temperature

$$p = n/2 \tag{29a}$$

$$n = 3F/(2\beta N) \tag{29b}$$

High Temperature ($P_3 \ll N$)

$$p = (\beta n^2)/P_3 \tag{30a}$$

$$n = (FP_3/N)^{\frac{1}{2}}/\beta. \tag{30b}$$

For the calculation shown plotted in Fig. 8, equations (28b) and (29b) apply; the ratio between n at low temperatures and n at intermediate temperatures is given by:

$$R = \{2\beta/(9\beta')\}^{\frac{1}{2}} \approx 47 \tag{31}$$

Direct excitation from imperfection levels

Thus far only the case of excitation across the and gap, with the creation of both free electrons and free holes, has been considered. Such transitions across the band gap constitute the major contribution to photoconductivity in "pure" materials and in many activated materials with low activator proportions.

Direct excitation from imperfection levels may be important, however, for certain ranges of excitation wavelength for materials with a high impurity concentration. We shall at this point, therefore, briefly consider the changes in the preceding picture which would be necessitated if excitation of centers 1 or centers 2 occurred directly.

If excitation of centers 2 occurred directly, and the assumption of negligible thermal freeing of holes from centers 2 is retained, the resulting phenomena would be the same as for a simple photoconductor with recombination centers of type 2 only. The phenomena associated with charge-carrier transfer between centers 1 and centers 2 would be absent. Similarly, if excitation of centers 1 occurred directly, and if thermal freeing of holes from centers 1 were negligible, results characteristic of a simple one-recombination-center photoconductor would be found; in this case the photosensitivity would be larger than the previously mentioned case by the ratio β/β'.

Some interesting changes in the results are obtained if the model previously described is assumed with the only change being a limitation of the excitation process to transitions between centers 1 and the conduction band. To describe this case, equations (2), (3) and (4) may be replaced by the following equations, respectively:

$$dn/dt = 0 = AFn_1 - n\sum_{i=1}^{3}\beta_{n_i}(N_i - n_i) + n_3P_3 \tag{32}$$

$$dp/dt = 0 = (N_1 - n_1)P_1 - p\sum_{i=1}^{3}\beta_{p_i}n_i \tag{33}$$

$$dn_1/dt = 0$$

$$= (N_1 - n_1)P_1 - \beta_{p_1}n_1p + \beta_{n_1}n(N_1 - n_1) - AFn_1 \tag{34}$$

where A is the probability of a photon being absorbed by centers 1, and where the terms concerned with direct free-electron-free-hole recombination have been omitted. No changes are required in equations (1), (5), and (6).

The unique feature resulting from this type of excitation is the saturation in n and p which occurs with increasing light intensity. For the simple

VOL. 1 1956/57

algebraic assumptions given in equation (19), it may be shown that saturation will occur at the following values of the parameters:

$$\left.\begin{aligned} p &= P_1/(2\beta) \\ n &= \{2NP_1/(3\beta)\}^{\frac{1}{2}} \\ n_1 &= \{2NP_1/(3\beta)\}^{\frac{1}{2}}/2 \\ AF &= 2N[\beta' + \{3\beta P_1/(2N)\}^{\frac{1}{2}}]. \end{aligned}\right\} \quad (35)$$

(These values for n, n_1, and AF are explicitly for the case where saturation occurs when $\beta(n+p) \gg P_3$.)

Direct excitation of centers 1 has the following effects:

(1) Supralinearity. If saturation occurs before the onset of supralinearity, supralinearity will not be found; if saturation occurs in the supralinear range, supralinearity will be terminated at that point; if saturation occurs above the supralinear range, supralinearity will occur as in the band-gap-excitation model.

(2) Rise time. For band-gap excitation, very slow rise of photocurrent with time, following an S-shape curve, is commonly found, as holes are transferred from centers 2 to centers 1. When centers 1 are excited directly, below saturation, a much more rapid rise will be found, and in fact overshoot on the rise can be very pronounced because of the slow thermal release of holes from centers 1 to be captured by centers 2. Above saturation, the rise will continue fast, but the overshoot will disappear.

(3) Decay time. For a given concentration of electrons, n, in the conduction band, the lifetime for direct excitation of centers 1 will in general be somewhat longer than for band-gap excitation, because a smaller concentration of holes will be in centers 2 and 3 for the same density of free electrons. This difference will increase as the light intensity is increased and saturation is approached.

Direct excitation of electrons from centers 3 or other trapping levels to the conduction band during the course of the decay will decrease the trap-determined decay time regardless of the type of initial excitation. The infra-red-quenching transition, from the valence band to centers 1, will also decrease the decay time by decreasing the free electron lifetime.

Application to luminescence

If either centers 1 or centers 2 are luminescence centers, the dependence of luminescence emission intensity on exciting light intensity and temperature can be determined from the previous calculations. If centers 1 are luminescent centers, the corresponding emission intensity, I_1, will be given by $\beta' n(N_1 - n_1)$; similarly if centers 2 are luminescent centers, I_2 will be given by $\beta n(N_2 - n_2)$. An illustrative calculation, with the same approximations as used for the determination of the variation of n with F in Fig. 3, shows that I_1 will vary supralinearily with F over a very wide range, as $F^{5/3}$ for $10^3 \leqslant F \leqslant 10^{11}$, and approximately as F^2 for $10^{11} \leqslant F \leqslant 10^{16}$. I_2, on the other hand varies almost exactly linearily with F over the entire range. If both I_1 and I_2 were present simultaneously, I_2 would be orders of magnitude larger than I_1 over the entire F range because $\beta \gg \beta'$.

DISCUSSION OF EXPERIMENTAL RESULTS

Photocurrent as a function of temperature and light intensity

In order to check the above theory and derive from it certain useful values for the parameters involved, measurements were made on the variation of photocurrent with incandescent light intensity (giving primarily band-gap excitation) at various temperatures for a single crystal of CdSe : I : Cu[13] ($0{\cdot}92 \times 0{\cdot}46 \times 0{\cdot}023$ mm^3). The data obtained are presented in two different ways: (1) as the variation of photocurrent with temperature for different intensities of excitation, in

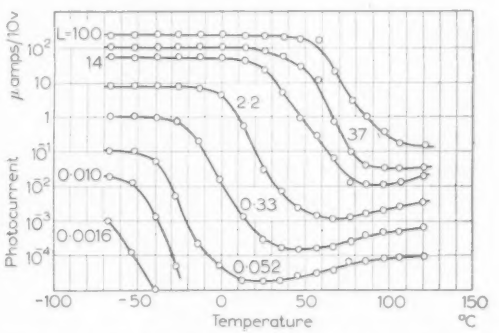


FIG. 9. Variation of photocurrent as a function of temperature for different relative intensities of excitation for a CdSe : I : Cu crystal.

Fig. 9, and (2) as the variation of photocurrent with light intensity for different temperatures, in Fig. 10. The experimental points are indicated on Fig. 9 only; the lines on Fig. 10 are drawn through the experimental points particularly to illustrate the break-points from near-linear to supralinear behavior. These data may be used to obtain values for the ratio β/β' of the centers 1, and for E_1, the energy difference between the levels 1 and the top of the valence band.

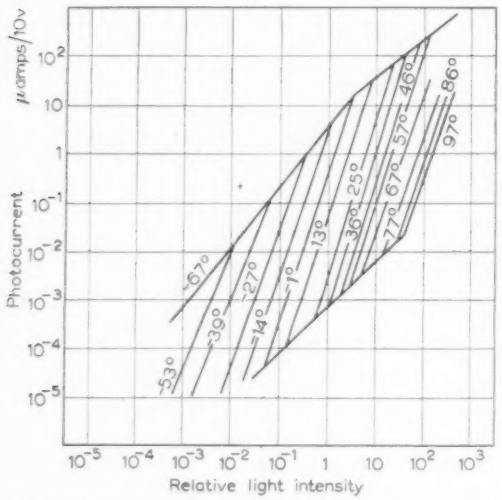


FIG. 10. Variation of photocurrent as a function of light intensity for different temperatures for a CdSe : I : Cu crystal.

An approximate estimate of the ratio β/β' can be obtained from the decrease in photosensitivity with temperature shown in Fig. 9, using equation (31) for the observed value of R of about a thousand. A value of $\beta/\beta' = 10^6-10^7$ is obtained from such an estimate.

Probably the best values for β/β' and E_1 can be obtained from the data of Figs. 9 and 10 by considering the values of n for the break from high to decreasing sensitivity, given by the condition that $\beta' n = P_1$. Considering the definition of P_1, this condition can be rewritten as:

$$\ln n = \ln(N_v\beta/\beta') - E_1/(kT). \qquad (36)$$

The log of the photocurrent is plotted against $1/T$ in Fig. 11 for all the values of the break from high to decreasing sensitivity obtained from both Fig. 9 and Fig. 10. From the slope of the straight line drawn through these points, a value of $E_1 = 0{\cdot}64$ eV is obtained. From the intercept of the line for $1/T = 0$, a value of $\beta/\beta' = 8\times10^5$ is obtained for a value of $N_v = 10^{19}/\text{cm}^3$.

The condition for the break to low sensitivity is given in equation (22) and may be rewritten as:

$$\ln n = \ln\{(N_2+N_3)N_v/N_1\} - E_1/kT. \qquad (37)$$

Thus if $\log n$ is plotted as a function of $1/T$ for the break from decreasing to low sensitivity, a straight line should be found with slope E_1 and intercept $(N_2+N_3)N_v/N_1$. The data have been plotted and are also shown in Fig. 11. The predicted slope of E_1 is found, and from the value of the intercept, again assuming $N_v = 10^{19}/\text{cm}^3$, a value of $(N_2+N_3)/N_1 = 2$ is indicated. Thus the assumption in other calculations of $N_1 \approx N_2 \gtrsim N_3$ seems reasonable.

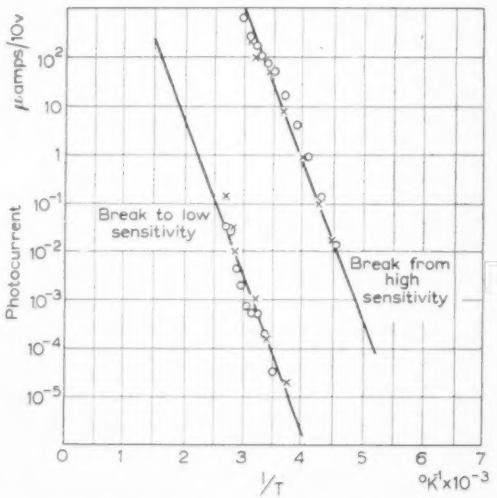


FIG. 11. Photocurrents at which the break from high sensitivity and the break to low sensitivity occur as a function of the reciprocal temperature. Circles are data from Fig. 10; crosses are data from Fig. 9.

Another way of analysing the data is given by the use of the fact that the break from high sensitivity corresponds to a location of the hole demarcation line for the centers 1 at the levels 1. Using equation (17) to calculate the location of the quasi-Fermi level for electrons from the photocurrent at the break, a plot of E_{F_n} as a function of T may be

made to test equation (14). The results are shown in Fig. 12. Although there is considerable scatter in the points, the best straight line drawn through the points is a fair representation of the data. Such a line has a slope which gives $\beta/\beta' = 10^5$ and an intercept for $T = 0$ of 0·59 eV which corresponds to E_1. A line drawn with slightly higher slope to better match the value of β/β' found in the previous method would also produce a larger value for E_1, in better agreement with the previous method.

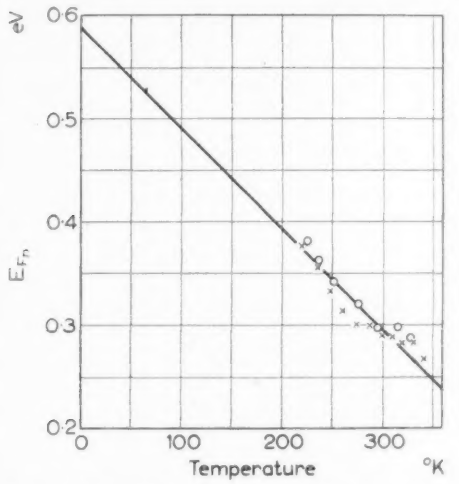


FIG. 12. Calculated values for the location of the quasi-Fermi level for electrons at the break from high sensitivity as a function of temperature. Circles are data from Fig. 9; crosses are data from Fig. 10.

Fig. 9 shows that the photosensitivity increases again slightly at high temperatures. Analysis of the data in terms of equation (30b) which predicts that $n \propto \exp\{-E_3/(2kT)\}$ for the three curves with $L = 2{\cdot}2$, 0·33, and 0·052, gives for the value of E_3, 0·44, 0·50, and 0·42 eV respectively. An average value of 0·45 eV agrees very well with the major trap peak in CdSe as obtained from thermally stimulated current[1, 3] (see Fig. 14).

Speed of response as a function of temperature and light intensity

Measurements of the photocurrent and the decay time (time for photocurrent to decrease to one-half) as a function of temperature for the CdSe : I : Cu crystal used in the previous measurements are compared with similar measurements for a CdS : Cl : Cu[9] crystal (0·80×0·70 ×0·017 mm³) in Fig. 13. It is found that the measured decay time has approximately the same temperature dependence as the photocurrent. The chief difference between CdS and CdSe is that the temperature for which an abrupt decrease in photocurrent and decay time sets in, is around 0°C for CdSe and around 100°C for CdS. This difference, associated with the greater energy difference between the levels of type 1 from the top of the valence band in CdS than in CdSe, is also the cause of the faster response of CdSe at room temperature.

The decay of photocurrent after the cessation of excitation can be analysed into the following components, listed in the order of decreasing decay rate:

(1) Direct recombination of free electrons with holes in recombination centers with a large capture cross-section for electrons, without any trapping process involved. (2) Emptying of shallow traps, the freed electrons recombining with holes in the large capture-cross-section centers. (3) (a) Transfer of holes from small capture-cross-section centers to large capture-cross-section centers, or (b) emptying of deep traps. The importance of each of these components in a measured decay curve will depend strongly on the light intensity used. For very high light intensities, for which the density of free electrons is much larger than the density of trapped electrons, the measured decay time will be equal to the true lifetime as determined by component (1). For very low light intensities, the measured decay time will be determined almost completely by the faster of the (3) components alone.

The measured decay time in CdSe is less than in CdS at room temperature for two reasons, the difference in decay times becoming more pronounced as the light intensity is lowered. First, the true lifetime of a free electron in CdSe at room temperature decreases with decreasing light intensity; i.e., because of the supralinearity found in CdSe at room temperature, the sensitivity and the lifetime decrease with decreasing light intensity. Second, transfer of holes from small capture-cross-section centers to large capture-cross-section centers occurs at a faster rate than the emptying of at least some of the deeper traps in CdSe (i.e., component (3a) dominates the decay rate),

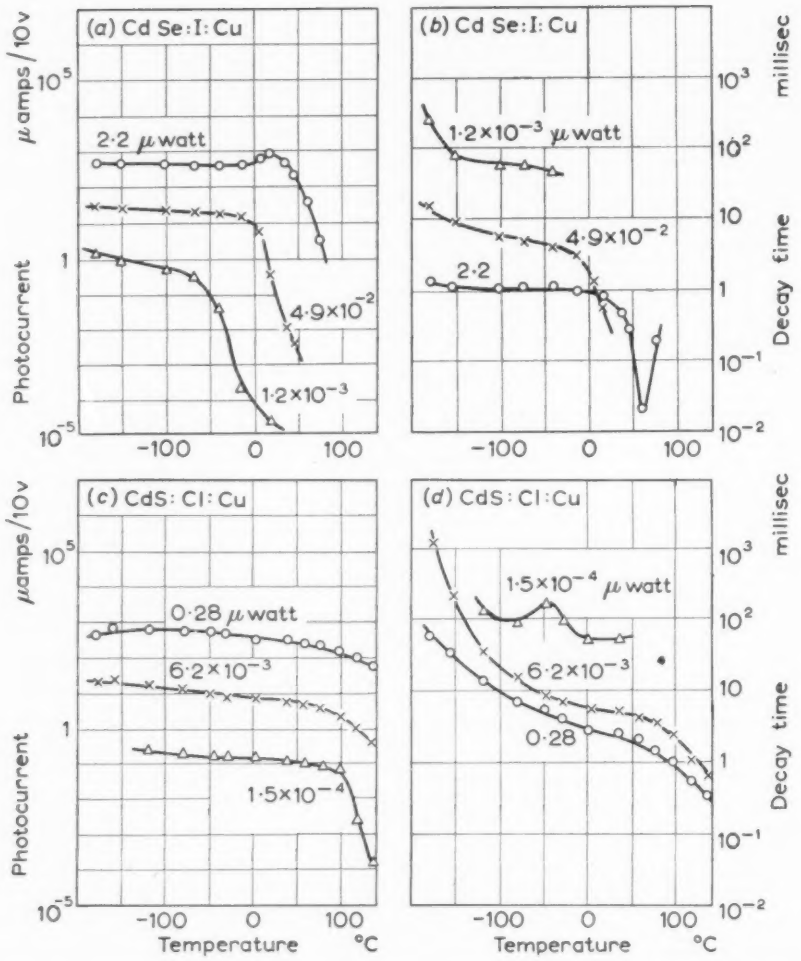


FIG. 13. Temperature dependence of photocurrent and measured decay time for various excitation intensities for a CdSe : I : Cu crystal and a CdS : Cl : Cu crystal.

whereas in CdS the transfer of holes occurs at a much slower rate than the emptying of deep traps (i.e., component (3b) dominates the decay rate). In order for the rate of thermal ionization of holes from centers 1 (lying 0·64 eV above the valence band) to be greater than the thermal ionization of electrons from those traps in CdSe lying 0·45 eV below the conduction band,

$$(N_1-n_1)S_{p_1}/(n_3S_{n_3})>2\times10^3 \qquad (38)$$

which is a reasonable relationship when it is considered that N_1 is probably larger than N_3 and that S_{p_1} (the cross-section for the capture of a free hole by a singly or doubly negative center) may well be considerably larger than S_{n_3}. Additional evidence for a relatively small value for S_{n_3} is given in the next section.

The above arguments have been made with the implicit assumption that the difference in the speed of response between CdS and CdSe does not depend on a significant difference in the density or distribution of trapping centers available.

A first glance at the thermally-stimulated current curves for the two crystals discussed in this paper, shown in Fig. 14, might indicate that there are far

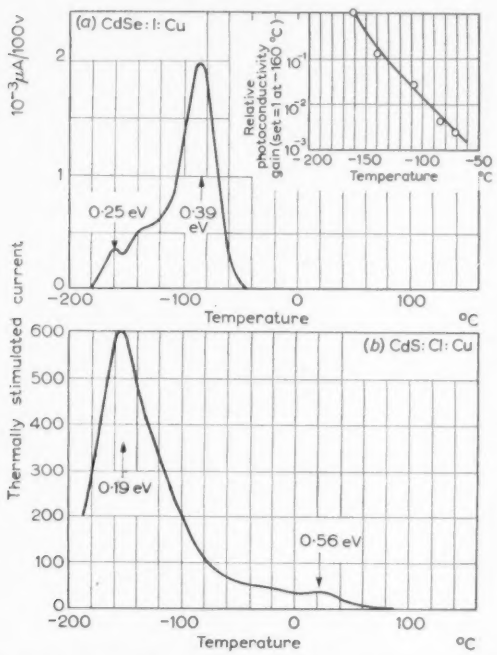


FIG. 14. Thermally stimulated current curves measured at 0·77°/sec for a CdSe : I : Cu crystal and a CdS : Cl : Cu crystal. Inset in (a) shows the estimated variation of photoconductivity gain (see text) as a function of temperature.

fewer traps in CdSe than in CdS. That the difference between CdSe and CdS is much less than this can be shown by a consideration of the actual conditions present as these curves were measured. All of the electrons freed from traps and contributing to the thermally stimulated current for the CdS crystal (except for temperatures above 50°C) have the lifetime of the sensitive CdS crystal. This lifetime will in fact increase to some extent with increasing temperature, as the electrons freed from traps are captured by the large cross-section centers and hence the number of empty large cross-section centers for subsequent electrons freed from traps becomes less. Hole transfer from the centers 1 to the centers 2 and 3 in CdS does not occur at an appreciable rate below temperatures of 50°C. To illustrate this fact, the calculated values of the quasi-Fermal level for electrons are listed on the curve; a quasi-Fermi level 0·60 eV from the conduction band in CdS corresponds to a location of the hole demarcation line for centers 1 at levels 1 (see eq. 14). On the other hand, hole transfer in CdSe from centers 1 to centers 2 and 3 sets in at about −160°C (the demarcation line for centers 1 being at levels 1 is indicated in CdSe by a quasi-Fermi level for electrons 0·3 eV from the conduction band). An estimate of the decrease in free-electron lifetime with increasing temperature because of this hole transfer process in CdSe is given in the insert of Fig. 14(a). The photoconductivity gain, determined from steady-state measurements under illumination, for a photocurrent equal to the thermally stimulated current at the temperature of the thermally stimulated current, is plotted as a function of temperature. The gain, and hence the approximate contribution to the current of an electron thermally freed from a trap, decreases by a factor of about 10^3 between −160° and −50°C because of the hole transfer process. This factor can be much larger than the increase in the lifetime of an electron freed from a trap because of the filling of recombination centers by previously freed electrons. The abrupt cutoff on the high-temperature side of the thermally stimulated current peak for CdSe is thus principally caused by the rapid decrease in the lifetime of electrons freed from traps with increasing temperature, and not by the trap density or distribution.

Rise and decay transients

Previous publications[1, 2, 3] have pointed out how the slow rise of photocurrent frequently observed in CdS and CdSe can be correlated with the sensitization of the crystal with time, the concentration of holes in centers 1 being increased, and the concentration of holes in centers 2 being decreased. That the reverse process during decay is a much quicker one (i.e., the transfer of holes from centers 1 to centers 2 after the cessation of excitation), as indicated in the discussion of the previous section, is shown also by the fact that the drop in photosensitivity with decreasing light intensity in the supralinear range occurs substantially simultaneously.

If the light intensity is suddenly lowered, either to zero or to a low value within the supralinear

range, the photosensitivity may indeed undershoot its long-time equilibrium value. Such undershoot will occur prominently only when the density of holes in centers 2 is equal to or larger than the density of free plus trapped electrons. It is, therefore, not found until measurements are made at a sufficiently high temperature so that most of the holes are located in centers 2 rather than centers 1 at the low light intensity. Under these conditions, the free and very shallow trapped electrons are drained off very rapidly by recombination, the rate of recombination ultimately falling below the rate of electron ejection from traps. When this occurs, the density of free electrons increases to the equilibrium value. An over-rise is found under similar circumstances when the light intensity is increased suddenly to a high value from the dark or from a low light intensity.

The necessary conditions for the existence of undershoot may be determined by considering the equation for the rate of change of the density of free electrons in the absence of excitation:

$$dn/dt = n_3 N_c S_{n_3} v \exp(-E_3/kT) - \\ -nvS_{n_2}(N_2-n_3) - nvS_{n_3}(N_2-n_2) \qquad (39)$$

The charge neutrality equation may be written as:

$$(n-n_0)+(n_3-n_{3_0}) = (n_{2_0}-n_2) \qquad (40)$$

where the subscript zero indicates the densities in the dark or at a low light intensity. It has been assumed that all the holes corresponding to the free and trapped electrons are located in the centers 2 at the start of the decay. This will still be substantially true in those regions where a rapid transfer of holes from centers 1 to centers 2 occurs practically simultaneously with the cessation of excitation.

The condition for a minimum in the variation of n with t can be found by setting equation (39) equal to zero and solving for n. If the calculated value for $n_{\min.}$ is less than n_0, then undershoot will occur.

It can be shown that $n_{\min.}$ is given by:

$$n_{\min.} = N_c \exp(-E_3/kT)(S_{n_3}/S_{n_2}) \times \\ \times\{1+(N_2-n_{2_0})/N_3\}^{-1} \qquad (41)$$

under the conditions that n_0 and n_{3_0} are both much less than N_2 and that n_3 is about equal to N_3 at the minimum in n.

We can express n_0:

$$n_0 = N_c \exp(-E_{F n_0}/kT). \qquad (42)$$

The condition for undershoot, $n_{\min.} < n_0$, will, therefore, be met if:

$$(S_{n_3}/S_{n_2}) < \exp\{-(E_{F n_0}-E_3)/kT\} \times \\ \times\{1+(N_2-n_{2_0})/N_3\}. \qquad (43)$$

If in the dark or at the low light intensity, the density of holes in centers 2 is substantially less than the density of centers 3, the last term in the above equation will drop out.

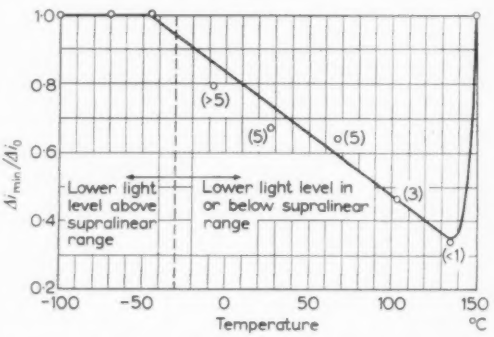


FIG. 15. Fractional undershoot of photocurrent for a "pure" CdSe crystal with a sudden decrease in light intensity, as a function of temperature. The small numbers in parentheses are the approximate times in minutes required for the photocurrent to increase from its minimum value to its low-light equilibrium value.

A plot of $n_{\min.}/n_0$, (actually $\Delta i_{\min.}/\Delta i_0$), as a function of temperature for a single crystal of "pure" CdSe, for a sudden change in light intensity by a factor of 300, is shown in Fig. 15. As long as the lower light intensity lies above the supralinear range, no undershoot is observed. When the lower light intensity lies in or below the supralinear range, undershoot is found which increases approximately linearly with temperature in magnitude, until a temperature is reached such that the condition of equation (43) no long applies. From the values of current and temperature for which the undershoot ceases, a value of S_{n_3}/S_{n_2} of about 10^3—10^4 can be estimated. That S_{n_3} is somewhat smaller than the other cross-sections under discussion is in agreement with the condition for rapid hole transfer of equation (38). The small numbers in parentheses below each point in Fig. 15 are the

approximate times in minutes required for the current to increase from $n_{min.}$ to n_0.

These results are similar to those reported by BOER[14] for CdS at elevated temperatures above 100°C, and attributed by him to possible photochemical reactions in CdS. That the effect should be found in CdS only above 100°C is in agreement with the previous mechanism being the cause of BOER's effects; most of the holes will not be concentrated in centers 2 in CdS until temperatures over 100°C are reached. The fact that the same type of phenomena can be obtained in CdSe below 0°C and can be explained by purely electronic processes, argues strongly against the necessity for photochemical reactions to achieve such effects.

CONCLUSIONS

A model of a photoconductor with two different types of recombination centers (one with a much smaller capture cross-section for electrons than the other), and trapping centers, is able to provide the mechanism for supralinear photoconductivity, for the temperature dependence of photosensitivity and speed of response, and hence by inference, for infra-red quenching (the raising of electrons from the valence band to levels 1 when the crystal is in a sensitive state).

A consideration of the onset and cessation of supralinearity, in the light of the theoretical concepts, leads to the determination from experimental data on CdSe of a ratio of 8×10^5 for the capture cross-section of the low-lying "sensitizing" centers for holes to that for electrons. It was also determined that the levels corresponding to these centers lie 0·64 eV above the valence band. The range of photocurrents and light intensities for which supralinearity is found was demonstrated to lie between (1) that condition for which the density of empty levels with small electron capture cross-section is comparable to the density of empty levels with a large electron capture cross-section, and (2) that condition for which the hole demarcation line for the small cross-section centers is located at the level corresponding to these centers. Thus it is shown that a range of supralinearity does not necessarily imply a distribution of "sensitizing" centers in energy.

Assuming the same value of the cross-section ratio of the "sensitizing" centers in CdS as in CdSe, results (according to eq. 14) in a location of the levels corresponding to these centers about 1·0 eV above the valence band. Such a value is in good agreement with the spectral distribution of infra-red quenching.[2]

Since the phenomena of supralinearity, and its associated characteristics, occur in crystals of CdSe and CdS to which no impurities have been added, it is indicated that the levels located at 0·64 eV above the valence band in CdSe and 1·0 eV above the valence band in CdS are associated with crystal defects. These crystal defects are probably cation vacancies with one or two trapped electrons, which are, therefore, singly or doubly negative with respect to the rest of the crystal. Supralinearity has been found in CdS at room temperature in samples with a high proportion of Cu impurity;[15] such results suggest that the Cu^+ level lies at about 0·6 eV above the valence band and also acts as a center with much larger cross-section for holes than for electrons. Measurements of transmission through CdS sintered layers with small proportions of Cu have revealed only a small absorption at 1·4 eV corresponding to the defect levels 1·0 eV above the valence band. Measurements of transmission and excitation spectrum for luminescence on CdS sintered layers with high Cu proportion* indicate high absorption and impurity-excitation of luminescence at about 1·8 eV, corresponding to a level 0·6 eV above the valence band.

The following tentative conclusions about band-gap-excited photoconductivity in CdS-type photoconductors can be made insofar as the model is valid.

(1) Almost all of the properties of interest to photoconductivity in these materials are associated with the presence of a set of small capture-cross-section imperfection levels, which may or may not act effectively as recombination centers, depending on the temperature and light intensity.

(2) In particular, the speed of response of CdSe is caused by the specific location of these defect levels, which location (for normal operation based at room temperature) has of necessity associated with it both marked temperature dependence of photosensitivity and speed of response, and supralinear photocurrent vs. light intensity variation.

* Obtained from H. A. KLASENS of Philips Co., Eindhoven; measurements by W. SPICER.

(3) The speed of response in all materials, especially at low excitation intensities, is limited ultimately by trapping phenomena.

(4) Materials fall into three general categories according to their properties at room temperature: (a) sensitive materials with slow response, small temperature dependence, and linear or sublinear photocurrent vs. light variation; (b) materials sensitive at high light intensities, with faster response than type (a) materials (particularly at low light intensities), large temperature dependence, and supralinear photocurrent vs. light variation; and (c) less sensitive or insensitive materials with intermediate or fast response, small temperature dependence, and linear or sublinear photocurrent vs. light variation.

Acknowledgements—The author wishes to express his appreciation to M. A. LAMPERT for many valuable discussions and for helpful suggestions in the writing of this paper. The treatment of "undershoot" phenomena resulted directly from such discussions. Talks with Dr. C. CALVO were also helpful in the early phases of the work.

REFERENCES

1. BUBE R. H. Paper in *Proceedings of the Conference on Photoconductivity, Atlantic City*, 1954, John Wiley and Son, in press.
2. BUBE R. H. *Phys. Rev.* **99,** 1105 (1955).
3. BUBE R. H. *Proc. I.R.E* **43,** 1836 (1955).
4. BUBE R. H. and LIND E. L. *Bull. Amer. Phys. Soc.* Ser. II, **1,** p. 110 (1956).
5. ROSE A. *RCA Rev.* **12,** 362 (1951).
6. ROSE A. Paper in *Proceedings of the Conference on Photoconductivity, Atlantic City*, 1954, John Wiley and Son, in press.
7. ROSE A. *Phys. Rev.* **97,** 322 (1955).
8. ROSE A. *Proc. I.R.E.* **43,** 1850 (1955).
9. DUBOC C. A. *Brit. J. Appl. Phys.* Supplement 4, p. S 107 (1955).
9a. DIRKSEN H. J. and MEMELINK O. W. *Appl. Sci. Res.* **4B,** 205 (1954).
10. SCHÖN M. *Physica* **20,** 930 (1954).
11. NEWMAN R., WOODBURY H. H., and TYLER W. W. *Phys. Rev.* **102,** 613 (1956).
12. TYLER W. W. and WOODBURY H. H. *Phys. Rev.* **102,** 647 (1956).
13. BUBE R. H. and THOMSEN S. M. *J. Chem. Phys.* **23,** 15 (1955).
14. BOER K. W. *Physica* **20,** 1103 (1954).
15. THOMSEN S. M. and BUBE R. H. *Rev. Sci. Instrum.* **26,** 664 (1955).

VOL. 1 1956/5

J. Phys. Chem. Solids. Pergamon Press 1957. Vol. 1. pp. 249–261.

BAND STRUCTURE OF INDIUM ANTIMONIDE

EVAN O. KANE

General Electric Research Laboratory, Schenectady, New York

(*Received* 23 August 1956)

Abstract—The band structure of InSb is calculated using the $k \cdot p$ perturbation approach and assuming that the conduction and valence band extrema are at $k = 0$. The small band gap requires an accurate treatment of conduction and valence band interactions while higher bands are treated by perturbation theory. A highly nonparabolic conduction band is found. The valence band is quite similar to germanium. Energy terms linear in k which cannot exist in germanium are estimated and found to be small, though possibly of importance at liquid-helium temperature. An absolute calculation of the fundamental optical absorption is made using the cyclotron resonance mass for n-type InSb. The agreement with experimental data for the fundamental absorption and its dependence on n-type impurity concentration is quite good. This evidence supports the assumptions made concerning the band structure.

1. INTRODUCTION AND CONCLUSIONS

THERE is now a considerable amount of experimental and theoretical information available concerning the properties of indium antimonide so that it has become possible to give a reasonably consistent picture of the band structure of this material. We will first present a summary of the information that is already known and then discuss what further properties can be deduced from the analysis presented in this paper.

Cyclotron resonance measurements[1] indicate that the conduction band minimum in indium antimonide is at $k = 0$ and is doubly degenerate* with an effective mass of $0{\cdot}013\ m_0$. Magnetoresistance measurements[2, 3] also support these conclusions.

The fundamental optical absorption measurements of FAN and GOBELI[4] show that the valence band energy at $k = 0$ is within a few hundredths of a volt of the thermal gap. These measurements were made at room temperature and gave an optical gap for direct transitions of 0·175 eV. ROBERTS and QUARRINGTON[5] have shown that the optical gap is strongly temperature dependent, having a value 0·23 eV at $T = 0$ K.

The following information about the conduction and valence bands in the neighborhood of $k = 0$ can be obtained from the work of DRESSELHAUS and PARMENTER[6] on the symmetry properties of the zinc-blende structure.

The energy of the valence band in the neighborhood of $k = 0$ will be proportional to k, rather than to k^2 as it is for germanium. The valence-band maxima will most probably occur in the (111) directions away from $k = 0$. For energies very near the maxima, the energy surfaces will be ellipsoids of revolution with their axes of revolution in the (111) directions. The fourfold degeneracy of the bands at $k = 0$ is completely removed for a general k direction. The maxima in the (111) direction are singly degenerate.

For larger values of k the linear terms will be dominated by terms proportional to k^2. These k^2 terms give energy bands entirely similar to germanium. The fourfold degeneracy at $k = 0$ is broken up into a doubly degenerate heavy-mass band and a doubly degenerate light-mass band. The double degeneracy is split by an amount proportional to k by the linear k terms. The double degeneracy of the conduction band is removed by terms proportional to k^3.

A number of experiments suggest the value

* This double degeneracy is due to spin.

0·18 m_0 as the "average"* valence band mass. Thermoelectric power measurements by H. WEISS[7] give 0·18 m_0 at room temperature. The binding energy of holes, 0·007 eV as determined by HROSTOWSKI *et al.*,[8] suggests that the "average" mass is less than the density of states mass for germanium, 0·35 m_0, since holes in germanium are bound by 0·01 eV. HROSTOWSKI found that an "average" mass of 0·18 m_0 gave a good fit to his data on the deionization of holes. Cyclotron resonance measurements[1] have shown the presence of mass peaks at 0·18 m_0 and $> 1{\cdot}2\, m_0$. These peaks were tentatively identified with holes.

The existence of light holes is suggested by the data of HROSTOWSKI *et al.*[8] on the variation of Hall coefficient with magnetic field. Their data are similar to the data of WILLARDSON *et al.*[9] which have been interpreted in terms of light and heavy holes.

There is a region of optical absorption of the order 100 cm^{-1} extending several hundredths of a volt beyond the edge for direct transitions. Two analyses of the absorption in this region have been made according to a method first proposed by MACFARLANE and ROBERTS.[10] The two analyses reach different conclusions concerning the band structure. BLOUNT *et al.* suggest that the valence band maximum is far from $k = 0$. POTTER stresses the importance of optical phonons and suggests that the valence band maximum is 0·015 eV above the energy at $k = 0$, but at a value of k much nearer zero than claimed by BLOUNT *et al.*

The effect of impurities on the band structure has been calculated by STERN and TALLEY.[11] They made the approximation that their impurities were periodically arranged so that they could calculate the band structure of metallic hydrogen with an appropriate effective mass and dielectric constant. The neglect of randomness is, of course, a strong objection to this procedure. Randomness can be taken into account by a perturbation approach as has been done by PARMENTER.[12] Perturbation theory cannot apply to bound or quasibound states but may be applicable to free states. The limitations of perturbation theory have not yet been adequately discussed for this problem. The perturbation calculation gives an appreciably smaller effect than the Stern and Talley approach.

The present paper enlarges upon the $k \cdot p$ approach of DRESSELHAUS and PARMENTER.[6] The spin-orbit interaction is also treated as a perturbation.[13] It has been assumed in this paper that because of the smallness of the energy gap at $k = 0$ it would be a good approximation to ascribe all of the unusually low effective mass of the conduction band to the mutual interaction of the conduction and valence bands.

A modification of the usual perturbation technique is used whereby the mutual interaction of the conduction and valence bands via the $k \cdot p$ interaction and the k-independent spin-orbit interaction are treated exactly. The effects of higher and lower bands are treated by perturbation theory.

For energies above $k = 0$ of the order of a fraction of the band gap the conduction band becomes non-parabolic. At this point straightforward perturbation theory breaks down but the present technique remains valid. Because of the small band gap, the non-parabolic nature of the conduction band will be important at high temperature or high electron concentration (degeneracy).

In diagonalizing the mutual interaction of the conduction and valence bands the sixfold degeneracy of the valence band is split into three twofold bands corresponding to the heavy-mass band, the light-mass band and the spin-orbit split-off band of germanium. The conduction band is also doubly degenerate. With part of the valence-band degeneracy thus removed, the interaction with higher bands is calculated by second-order perturbation theory. The second-order terms involving the $k \cdot p$ interaction alone shift the bands by an amount proportional to k^2. The remaining twofold degeneracy of the bands is removed by terms proportional to k^3. The twofold degeneracy of the valence band is also removed by second-order terms involving the $k \cdot p$ interaction and the k-independent spin-orbit interaction, giving a splitting proportional to k. These second-order terms are larger than the first-order linear k terms

* For the case of non-spherical energy surfaces the concept of an "average mass" has no precise significance since it will in general be different for every different experiment. The above summary of the experimental information for the valence band uses the word "average" loosely, but it should be understood that a more detailed analysis of each experiment in terms of a definite model is necessary in order to derive maximum information.

VOL. 1 1956/5

considered by DRESSELHAUS[6] which resulted from the k-dependent spin-orbit interaction.

The above procedure results in simpler formulae for the valence bands than were obtained by DRESSELHAUS. The simplicity results from having removed part of the degeneracy before applying perturbation theory. Before perturbation theory is applied the bands are spherical. The accuracy of the simpler expressions is less than that of the more complicated expressions. Considering the k^2 terms alone, the present procedure will be accurate if the mutual interaction energy of conduction and valence bands is large compared to the interaction energy with higher bands. This is the case in InSb. The present method always breaks down for energies very close to $k = 0$ where the linear k terms dominate the k^2 terms used to remove part of the initial valence band degeneracy.

The spin-orbit splitting of the valence band is not known but a rough estimate from atomic spectra puts it at about 0·9 eV. According to theory the light mass should be 0·015 m_0 with a small uncertainty due to the unknown spin-orbit splitting. The effect of higher bands on the light mass value should also be small. The heavy mass, however, is determined entirely by higher bands, receiving no contribution from the conduction band interaction. It is reasonable to assume tentatively that the 0·18 m_0 average valence band mass is the average value for the heavy mass band. No estimate of the warping of the heavy mass band has yet been obtained. There appears to be no place in the present theory for the $> 1{\cdot}2\ m_0$ peak seen in cyclotron resonance.

The magnitude of the linear k terms is estimated very roughly in the present paper. The linear k terms cause the maximum in the valence band to be $\sim 10^{-4}$ eV above the energy at $k = 0$ for a value of $k \sim 0{\cdot}3$ per cent of the way to the zone boundary in the (111) direction. For energies much greater than 10^{-4} eV the k^2 terms are dominant and the valence band of indium antimonide resembles that of germanium. The above estimate of the magnitude of the linear k terms is very close to the value given by DRESSELHAUS.[6] However the present estimate refers to a second order process involving the $k \cdot p$ interaction and the k-independent spin-orbit interaction. These terms are actually larger than the first-order terms from the k-dependent spin-orbit interaction which were somewhat overestimated by DRESSELHAUS.

The fundamental optical absorption is calculated in the approximation of direct transitions. The optical matrix element is determined from the electron cyclotron mass so that no arbitrary constants are involved. The agreement with the experimental data of FAN and GOBELI[4] supports the major assumptions made in this paper.

The effect of impurities in shifting the optical absorption edge has been computed. The effects of a change in band structure are negligible using PARMENTER'S[12] perturbation theory calculation. The agreement with the experimental data of FAN and GOBELI[4] for 10^{18} impurities per c.c. is satisfactory.

The indirect transitions analysis of POTTER[10] would agree with the picture of the band structure presented here if it were assumed that the linear k terms were a good deal larger than the theoretical estimate. At the present time the theoretical analysis of indirect transitions does not appear sufficiently straightforward to warrant drawing strong conclusions.

The work of HROSTOWSKI *et al.*[8] at 77°K appears to suggest the existence of light holes. Since the mass ratio of light and heavy holes is similar in germanium and indium antimonide the variation of Hall coefficient with magnetic field should be similar. If the size of the linear k terms were as great as POTTER'S analysis requires, the band structure at energies appropriate to 77°K would be so unlike germanium that we could not speak of "light holes".

2. BAND STRUCTURE

General.

The point of view in the present calculations is similar to the point of view taken in an earlier paper[13] in discussing the valence band structure of germanium. The difference is that here the interaction of the valence bands and the conduction band via the $k \cdot p$ interaction and the k-independent spin-orbit interaction is taken into account exactly. The effect of higher bands is taken into account by second order perturbation theory. This refinement was necessary because the band gap is so small in indium antimonide that the parabolic bands given by a straight perturbation treatment would be an inaccurate

approximation to use in analysing many experiments of practical interest.

The Schroedinger equation for the cell-periodic functions $u_k(r)$ is

$$\{(p^2/2m)+V+(\hbar/m)\boldsymbol{k}\cdot\boldsymbol{p}+(\hbar/4m^2c^2)[\nabla V\times\boldsymbol{p}]\cdot\sigma+ +(\hbar^2/4m^2c^2)[\nabla V\times\boldsymbol{k}]\cdot\sigma\}u_k(r)=E'_k u_k(r) \tag{1}$$

$$E'_k=E_k-(\hbar^2/2m)k^2 \tag{2}$$

where E_k is the energy of a state with wave vector $\boldsymbol{k}$. We assume we know the solution of the equation

$$\{(p^2/2m)+V\}U_i=E_iU_i \tag{3}$$

and use the complete set of U_i as a basis for a representation. Group theory gives us the symmetry properties of the functions U_i. The functions corresponding to the conduction band are singly degenerate (doubly degenerate counting spin) and will be denoted by $S\uparrow$ and $S\downarrow$. The symbol $\uparrow$ means spin up and $\downarrow$ means spin down. The designation S refers to the fact that the functions have the symmetry properties of s functions under the operations of the tetrahedral group.[6] The functions S belong to symmetry type Γ_1 in the notation of reference 6. The valence band functions are triply degenerate (sixfold with spin) and will be written $X\uparrow$, $Y\uparrow$, $Z\uparrow$, $X\downarrow$, $Y\downarrow$, $Z\downarrow$. Again the notation serves to indicate that the functions have the symmetry properties of the atomic p-functions x, y, z under the operations of the tetrahedral group. These functions have symmetry Γ_4 in the notation of reference 6.

The third term in eq. (1) will be referred to as the $\boldsymbol{k}\cdot\boldsymbol{p}$ interaction, the fourth term as the k-independent or atomic-like spin-orbit interaction, and the fifth term as the k-dependent spin-orbit interaction. For the present we shall neglect the fifth term. Later we treat it as a perturbation and estimate its size. It turns out to be very small because the crystal momentum $\hbar\,k$ is very small compared to the atomic momentum $\boldsymbol{p}$ in the far interior of the atom where most of the spin-orbit interaction occurs.

The fifth term gives an energy linear in k in the valence band. Energies linear in k also result from second order perturbation terms containing $k\cdot p$ and the k-independent spin-orbit interaction. These second-order terms are actually larger than the first-order terms of the k-dependent spin-orbit interaction. The magnitudes of these terms will be discussed later.

The linear k terms are responsible for the fact that the valence band of indium antimonide is more complex than that of germanium.[6] We include these terms by perturbation theory. This procedure is only valid for energies such that the k^2 terms which remove the valence band degeneracy are large compared to the k terms. For lower energies the more complicated expressions given by DRESSELHAUS[6] must be used.

Valence and Conduction Band Interactions

We take the k vector in the z direction and consider the Hamiltonian corresponding to the first four terms of eq. (1). The mutual interaction of the conduction and valence bands leaves the bands doubly degenerate. We take as a basis $|iS\downarrow\rangle$, $|(X-iY)\uparrow/\sqrt{2}\rangle$, $|Z\downarrow\rangle$, $|(X+iY)\uparrow/\sqrt{2}\rangle$, $|iS\uparrow\rangle$, $|-(X+iY)\downarrow/\sqrt{2}\rangle$, $|Z\uparrow\rangle$, $|(X-iY)\downarrow/\sqrt{2}\rangle$. The first four functions are respectively degenerate with the last four. The 8×8 interaction matrix may be written

$$\begin{bmatrix} H & 0 \\ 0 & H \end{bmatrix} \tag{4}$$

$$H=\begin{bmatrix} E_s & 0 & kP & 0 \\ 0 & E_p-\Delta/3 & \sqrt{2}\Delta/3 & 0 \\ kP & \sqrt{2}\Delta/3 & E_p & 0 \\ 0 & 0 & 0 & E_p+\Delta/3 \end{bmatrix}.$$

The positive constant Δ is the spin-orbit splitting of the valence band. The real quantity P is defined by

$$P=-i(\hbar/m)\langle S|p_z|Z\rangle$$

$$\Delta=\frac{3\hbar i}{4m^2c^2}\left\langle X\left|\frac{\partial V}{\partial x}p_y-\frac{\partial V}{\partial y}p_x\right|Y\right\rangle. \tag{5}$$

E_s and E_p refer to the eigenvalues of the Hamiltonian of eq. (3). E_s corresponds to the conduction band and E_p to the valence band. Symmetry properties have been used. If the k vector is not in the z direction, the Hamiltonian is more complicated but it can be transformed to the form of eq. (4) by a rotation of the basis functions.

VOL. 1 1956/57

$$\begin{bmatrix} \uparrow' \\ \downarrow' \end{bmatrix} = \begin{bmatrix} e^{-i\phi/2}\cos\theta/2 & e^{i\phi/2}\sin\theta/2 \\ -e^{-i\phi/2}\sin\theta/2 & e^{i\phi/2}\cos\theta/2 \end{bmatrix} \begin{bmatrix} \uparrow \\ \downarrow \end{bmatrix} \tag{6}$$

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{7}$$

$$S' = S. \tag{8}$$

The angles θ, ϕ are the usual polar angles of the k vector referred to the crystal symmetry axes x, y, and z with θ measured from z and ϕ measured from x. This transformation would be obvious if the functions X, Y, Z transformed like the spherical harmonics x, y, z under the full spherical group rather than just under the tetrahedral group.

The four double roots of the secular equation resulting from eq. (4) can be written

$$E' = 0 \tag{9}$$

$$E'(E'-E_G)(E'+\Delta)-k^2P^2(E'+2\Delta/3) = 0 \tag{10}$$

where

$$\begin{aligned} E_s &= E_G \\ E_p &= -\Delta/3 \end{aligned} \tag{11}$$

E' is defined in eq. (2). E_G is the band gap at $k = 0$.

For small values of k^2 the solutions of eq. (9) and (10) give parabolic bands.

$$\begin{aligned} E_c &= E_G+\frac{\hbar^2k^2}{2m}+\frac{P^2k^2}{3}\left(\frac{2}{E_G}+\frac{1}{E_G+\Delta}\right) \\ E_{v1} &= \hbar^2k^2/2m \\ E_{v2} &= \frac{\hbar^2k^2}{2m}-\frac{2P^2k^2}{3E_G} \\ E_{v3} &= -\Delta+\frac{\hbar^2k^2}{2m}-\frac{P^2k^2}{3(E_G+\Delta)}. \end{aligned} \tag{12}$$

The above equations give the expressions for the conduction band, the heavy-mass band, the light-mass band, and the split-off band in that order. The equations are, of course, not complete because the effects of higher bands have not yet been included. The expressions are probably fairly good for all bands except the heavy-mass band E_{v1}.

A different approximation can be made in simplifying the solutions of eq. (10), namely that $\Delta \gg kP, E_G$. The solutions then become

$$\begin{aligned} E_c &= \frac{\hbar^2k^2}{2m}+\frac{E_G+(E_G^2+8P^2k^2/3)^{\frac{1}{2}}}{2} \\ E_{v1} &= \hbar^2k^2/2m \\ E_{v2} &= \frac{\hbar^2k^2}{2m}+\frac{E_G-(E_G^2+8P^2k^2/3)^{\frac{1}{2}}}{2} \\ E_{v3} &= -\Delta+\frac{\hbar^2k^2}{2m}-\frac{P^2k^2}{3(E_G+\Delta)}. \end{aligned} \tag{13}$$

These equations exhibit simply the nonparabolic nature of the bands E_c and E_{v2} when $|(Pk)| \gg E_G$.

The doubly degenerate wave functions which result from the diagonalization of the Hamiltonian of eq. (4) may be written

$$\begin{aligned} \phi_{i\alpha} &= a_i[iS\downarrow]'+b_i[(X-iY)\uparrow/\sqrt{2}]'+c_i[Z\downarrow]' \\ \phi_{i\beta} &= a_i[iS\uparrow]'+b_i[-(X+iY)\downarrow/\sqrt{2}]'+c_i[Z\uparrow]' \\ \phi_{v1\alpha} &= [(X+iY)\uparrow]'/\sqrt{2} \\ \phi_{v1\beta} &= [(X-iY)\downarrow]'/\sqrt{2}. \end{aligned} \tag{14}$$

a_i, b_i, c_i are real coefficients; the index i refers to the bands E_c, E_{v2}, E_{v3}. The primes denote that the wave functions for a general k direction are rotated according to eqs. (6) and (7). The general expressions for a_i, b_i, c_i may be written

$$\begin{aligned} a_i &= kP(E_i'+2\Delta/3)/N \\ b_i &= (\sqrt{2}\Delta/3)(E_i'-E_G)/N \\ c_i &= (E_i'-E_G)(E_i'+2\Delta/3)/N \end{aligned} \tag{15}$$

where N is a normalizing factor equal to the square root of the sum of the squares of the numerators. The E_i' are the solutions of eq. (10). For small k eqs. (15) become

$$\begin{aligned} a_c &= 1,\ b_c = c_c = 0 \\ a_{v2} &= 0,\ b_{v2} = (1/3)^{\frac{1}{2}},\ c_{v2} = (2/3)^{\frac{1}{2}} \\ a_{v3} &= 0,\ b_{v3} = (2/3)^{\frac{1}{2}},\ c_{v3} = -(1/3)^{\frac{1}{2}}. \end{aligned} \tag{16}$$

In the approximation $\Delta \gg kP$, E_G the eqs. (15) become

$$a_c=\left(\frac{\eta+E_G}{2\eta}\right)^{\frac{1}{2}};\ b_c=\left(\frac{\eta-E_G}{6\eta}\right)^{\frac{1}{2}};$$

$$c_c=\left(\frac{\eta-E_G}{3\eta}\right)^{\frac{1}{2}}$$

$$a_{v2}=-\left(\frac{\eta-E_G}{2\eta}\right)^{\frac{1}{2}};\ b_{v2}=\left(\frac{\eta+E_G}{6\eta}\right)^{\frac{1}{2}};$$

$$c_{v2}=\left(\frac{\eta+E_G}{3\eta}\right)^{\frac{1}{2}} \tag{17}$$

$$a_{v3}=0;\ b_{v3}=(2/3)^{\frac{1}{2}};\ c_{v3}=-(1/3)^{\frac{1}{2}}$$

$$\eta=[E_G{}^2+8k^2P^2/3]^{\frac{1}{2}}.$$

Perturbations From Other Bands

The perturbing effects of higher and lower bands on the conduction and valence bands can be taken into account by second order perturbation theory. We take as our representation the complete set of functions U_i of the Hamiltonian operator of eq. (3) except that for the valence and conduction bands we use the wave functions of eq. (14).

We define a set of constants analogous to those used by DRESSELHAUS *et al.*[14]

$$A=\frac{\hbar^2}{m^2}\sum_{n,l}\frac{|\langle X|p_x|\Gamma_1(nl)\rangle|^2}{E_0-E_n}$$

$$B=\frac{\hbar^2}{m^2}\sum_{n,l}\frac{|\langle X|p_x|\Gamma_3(nl)\rangle|^2}{E_0-E_n}$$

$$C=\frac{\hbar^2}{m^2}\sum_{n,l}\frac{|\langle X|p_y|\Gamma_4(nl)\rangle|^2}{E_0-E_n}$$

$$D=\frac{\hbar^2}{m^2}\sum_{n,l}\frac{|\langle X|p_y|\Gamma_5(nl)\rangle|^2}{E_0-E_n}$$

$$F=\frac{\hbar^2}{m^2}\sum_{n,l}\frac{|\langle S|p_x|\Gamma_4(nl)\rangle|^2}{E_0-E_n}$$

$$G=\frac{\hbar^2}{m^2}\sum_{n,l}\frac{\langle S|p_x|\Gamma_4(nl)\rangle\langle\Gamma_4(nl)|p_y|Z\rangle}{E_0-E_n}. \tag{18}$$

The symbol $\Gamma_i\ (nl)$ refers to the lth function of the nth band having symmetry type Γ_i in the notation of reference 6. The summation, of course, does not extend over the conduction and valence bands. The energy E_0 is the valence band energy at $k=0$. More accurate expressions for the perturbation energies would result if we used the appropriate E_i given by eqs. (9) and (10) in place of E_0. The number of unknown parameters would be much larger if this were done.

In determining the effects of higher bands on the wave functions of eq. (14) we consider second-order perturbation matrix elements between degenerate wave functions $\phi_{i\alpha}$, $\phi_{i\beta}$ but not between functions ϕ_i, ϕ_j whose degeneracy has been removed by the interactions already discussed. In terms of the above constants the perturbing matrix elements can be written

$$\begin{aligned}H_{\alpha\alpha}(i)&=H_{\beta\beta}(i)\\&=a_i{}^2Fk^2+b_i{}^2\left\{(C+D)k^2+\left(\frac{3B}{2}-2C\right)\frac{k_x{}^2k_y{}^2+k_x{}^2k_z{}^2+k_y{}^2k_z{}^2}{k^2}\right\}+\\&\quad+c_i{}^2\left\{(A+B)k^2+(-3B+4C)\left(\frac{k_x{}^2k_y{}^2+k_x{}^2k_z{}^2+k_y{}^2k_z{}^2}{k^2}\right)\right\}\end{aligned} \tag{19}$$

$$H_{\alpha\beta}(i)=2\sqrt{2}a_ib_iG\left\{\frac{kk_z(k_y{}^2-k_x{}^2)+ik_xk_y(3k_z{}^2-k^2)}{k(k_x{}^2+k_y{}^2)^{\frac{1}{2}}}\right\}$$

$$|H_{\alpha\beta}(i)|=\frac{2\sqrt{2}|a_ib_iG|}{k}\{(k_x{}^2+k_y{}^2)(k_x{}^2+k_z{}^2)(k_y{}^2+k_z{}^2)-8k_x{}^2k_y{}^2k_z{}^2\}^{\frac{1}{2}}. \tag{20}$$

VOL. 1 1956/57

The index i refers to the bands E_c, E_{v2} and E_{v3}. The perturbation energy is then

$$H_{\alpha\alpha}(i) \pm H_{\alpha\beta}(i) \tag{21}$$

which is to be added to the energy determined from the cubic equation (10). The term $H_{\alpha\beta}$ removes the double degeneracy of the bands except for the (100) and (111) directions.

For the heavy-mass band E_{v1}

$$H_{\alpha\alpha} = (C+D)k^2 + \left(\frac{3B}{2} - 2C\right)\frac{k_x^2 k_y^2 + k_x^2 k_z^2 + k_y^2 k_z^2}{k^2}$$

$$H_{\alpha\beta} = 0. \tag{22}$$

The four constants A, B, C, D occur in three combinations which correspond to the constants L, M, N of reference 14. The constant D arises from spherical harmonics f and higher, so that it probably will be quite small.

The expression for the energy of the heavy mass band E_{v1} is simpler than the expression given by DRESSELHAUS *et al.*[14] for germanium. The reason for the simplification is that the degeneracy between the low-mass and high-mass valence bands at $k = 0$ is removed by the conduction band interaction alone in the present treatment. In the more accurate treatment of reference 14, all bands act to remove this degeneracy. The terms neglected in the present treatment are of the order of $E_G A/P^2$, $E_G B/P^2$, $E_G C/P^2$, $E_G D/P^2$ where P is the matrix element of eq. (5). The expressions given by the present treatment are identical with the expressions of reference 14 for the (100) and (111) directions.

Linear Terms

Energies linear in k are possible in the valence bands E_{v1} and E_{v2}. These linear terms act to split the double degeneracy of the bands E_{v1} and E_{v2}. No such terms are possible in the conduction band. Linear terms are altogether absent in germanium because of inversion symmetry. In the absence of spin-orbit interaction linear k terms are absent by time reversal symmetry even without inversion symmetry. These facts have been pointed out in reference 6.

The linear k terms are represented by an interaction Hamiltonian which has the form

$$H = K\begin{bmatrix} 0 & \sqrt{3}(k_x - ik_y) & k_x + ik_y & -2k_z \\ \sqrt{3}(k_x + ik_y) & 0 & -2k_z & -(k_x - ik_y) \\ k_x - ik_y & -2k_z & 0 & \sqrt{3}(k_x + ik_y) \\ -2k_z & -(k_x + ik_y) & \sqrt{3}(k_x - ik_y) & 0 \end{bmatrix} \tag{23}$$

with respect to a set of basis functions $|(X+iY)\uparrow/\sqrt{2}\rangle$, $|(X-iY)\downarrow/\sqrt{2}\rangle$, $|(2/3)^{1/2}Z\uparrow - (1/6)^{1/2}(X+iY)\downarrow\rangle$, $|(2/3)^{1/2}Z\downarrow + (1/6)^{1/2}(X-iY)\uparrow\rangle$.

Any set of Γ_8 functions may be shown to have a linear k Hamiltonian of this form if a representation is chosen having the same tetrahedral symmetry as the above basis functions.

The constant K arises from matrix elements of three types.

$$K_1 = -\frac{\hbar^2}{4\sqrt{3}m^2c^2}\left\langle X\left|\frac{\partial V}{\partial y}\right|Z\right\rangle \tag{24}$$

$$K_2 = -\frac{\hbar^2}{4\sqrt{3}m^3c^2}\sum_{n,j}\frac{\left\langle X\left|p_x\right|\Gamma_3(n,j)\right\rangle\left\langle\Gamma_3(n,j)\left|\frac{\partial V}{\partial z}p_x - \frac{\partial V}{\partial x}p_z\right|Y\right\rangle}{E_0 - E_n}$$

$$K_3 = \frac{\hbar^2}{2\sqrt{3}m^3c^2}\sum_{n,j}\frac{\left\langle X\left|p_y\right|\Gamma_5(n,j)\right\rangle\left\langle\Gamma_5(n,j)\left|\frac{\partial V}{\partial z}p_x - \frac{\partial V}{\partial x}p_z\right|X\right\rangle}{E_0 - E_n}.$$

The constant K_1 results from first-order perturbation theory applied to the k-dependent spin-orbit interaction, the fifth term in eq. (1). The constants K_2 and K_3 result from second-order perturbation theory applied to the $k \cdot p$ term and the k-independent spin-orbit interaction term in eq. (1). The constants K_2 and K_3 are actually much larger than K_1 because the momenta which make most of the contribution to the spin-orbit splitting are very large compared to $\hbar k$.

For small energies, the linear k terms of eq. (23) will dominate all the other terms which have been considered, since these terms vary like k^2 or higher powers of k. DRESSELHAUS[6] has discussed the band structure appropriate to the valence band taking into account linear k terms and second order perturbation $k \cdot p$ terms. The constants L, M, N which he uses are related to the constants defined in this paper as follows:

$$\begin{aligned} L &= A+B-P^2/E_G \\ M &= C+D \\ N &= A-(P^2/E_G)-(B/2)+C-D. \end{aligned} \tag{25}$$

For large energies, the k^2 and higher terms will be dominant and the terms of eq. (23) can be treated as perturbations.

The matrix elements of the linear k interaction between the heavy-mass band wave functions of eq. (14) are

$$\begin{aligned} H_{v1\alpha\alpha} &= H_{v1\beta\beta} = 0 \\ H_{v1\alpha\beta} &= 3\sqrt{3}K\left[\frac{kk_z(k_x^2-k_y^2)+ik_xk_y(k^2+k_z^2)}{k^2(k_x^2+k_y^2)^{\frac{1}{2}}}\right]. \end{aligned} \tag{26}$$

The perturbation energy is given by

$$\delta E_{v1} = \pm 3\sqrt{3}(K/k^2)\times \times[(k_x^2+k_y^2)(k_x^2+k_z^2)(k_y^2+k_z^2)]^{\frac{1}{2}}. \tag{27}$$

The bands are split in all but the (100) direction.

The matrix elements of the linear k interaction between the light-mass bands are

$$\begin{aligned} H_{v2\alpha\alpha} &= H_{v2\beta\beta} = 0 \\ H_{v2\alpha\beta} &= \\ &-\sqrt{3}K\left[\frac{kk_z(k_x^2-k_y^2)+ik_xk_y(k^2-3k_z^2)}{k^2(k_x^2+k_y^2)^{\frac{1}{2}}}\right] \end{aligned} \tag{28}$$

referred to the basis functions

$$\begin{aligned} \Phi_{v2\alpha} &= \left[\frac{(X-iY)}{6^{\frac{1}{2}}}\Big|\uparrow + \left(\frac{2}{3}\right)^{\frac{1}{2}} Z\Big|\downarrow\right]' \\ \Phi_{v2\beta} &= \left[-\frac{(X+iY)}{6^{\frac{1}{2}}}\Big|\downarrow + \left(\frac{2}{3}\right)^{\frac{1}{2}} Z\Big|\uparrow\right]'. \end{aligned} \tag{29}$$

The primes denote rotation by the operators of eqs. (6) and (7). The form of the light-mass band wave functions is less general than the form given in eq. (14) and is valid only in the neighborhood of $k = 0$. The reason for the restriction is that it is only for this simpler form that the three interactions of eq. (24) give the same result. The perturbation energy is

$$\delta E_{v2} = \pm\sqrt{3}(K/k^2)\times \times[(k_x^2+k_y^2)(k_x^2+k_z^2)(k_y^2+k_z^2)-8k_x^2k_y^2k_z^2]^{\frac{1}{2}}. \tag{30}$$

The bands are split in all but the (100) and (111) directions.

The linear terms give a maximum in the valence band away from $k = 0$ in the (111) direction. This maximum is associated with the heavy mass band E_{v1}. The simple expression (27) may be of semiquantitative value in discussing the energy surfaces in the neighborhood of the valence band maximum because the low mass band energy at this point in k space is large compared to the linear k-term energy, as the approximation leading to (27) requires. An improved expression could be obtained by doing second order perturbation theory with the linear k interaction.

The energy surfaces in the neighborhood of the valence band maxima are ellipsoids of revolution with the axis of revolution in the (111) direction. The maxima occur at a value of k given by

$$k_0 = \left|\frac{2\sqrt{2}K}{(\hbar^2/m)+B+(2C/3)+2D}\right|. \tag{31}$$

The energy of the maxima relative to $k = 0$ is

$$E_0 = \left|\frac{4K^2}{(\hbar^2/m)+B+(2C/3)+2D}\right|. \tag{32}$$

The longitudinal and transverse masses are given by the expressions

$$\begin{aligned} (1/m_l) &= (1/\hbar^2)[(\hbar^2/m)+B+(2C/3)+2D] \\ (1/m_t) &= (1/\hbar^2)[(11C/3)+3D-(B/2)+3\hbar^2/2m]. \end{aligned} \tag{33}$$

The transverse mass was derived using eq. (27) and eq. (22) hence it is only approximate. The other information can be derived from the expression for the energy in the (111) direction given by DRESSELHAUS.[6] See eq. (13) of this reference.

3. ESTIMATES OF BAND STRUCTURE CONSTANTS

The best determination of the band gap, E_G, appears to be the fundamental optical-absorption edge data of ROBERTS and QUARRINGTON.[5] Their data indicate a highly temperature-dependent edge

VOL. 1 1956/5

which extrapolates to a value $E_G = 0{\cdot}23$ eV at $T = 0°$K.

The spin-orbit splitting of the valence band is an unknown quantity which no experiment has indicated directly in the case of indium antimonide. The spin-orbit splittings of atomic p functions in indium and antimony are respectively 0·28 eV and ~0·8 eV from atomic spectra. As ELLIOTT[15] has pointed out, the atomic splittings are useful because the main spin-orbit interaction occurs so deep in the atom that the principal effect of the solid is to alter the normalization and to contaminate the wave function with higher spherical harmonics. The only reliable evidence of the magnitude of this effect comes from germanium where the one-electron atomic splitting is 0·20 eV and the solid splitting is 0·29 eV as shown by free carrier absorption.[16] The average ionicity of the valence wave functions in InSb has been estimated to be 0·35 by analysis of the reflectivity in the region of the lattice absorption.[17]

For want of a better estimate we assume that the valence electron at $k=0$ spends 35 per cent of its time on indium and 65 per cent of its time on antimony though this is not necessarily implied by the reflectivity analysis. If we make this assumption about the wave function and correct for normalization by the factor appropriate to germanium we get an estimated spin-orbit splitting of the order of 0·9 eV for indium antimonide.

The effective mass in the conduction band is 0·013 m_0 as given by cyclotron resonance measurements.[1] From the standpoint of the $k \cdot p$ approach it seems reasonable to attribute this unusually low mass to the strong repulsion of the closely adjacent conduction and valence bands. If we assume that the conduction-band–valence-band interaction is much larger than all other interactions responsible for the conduction electron mass we can determine the absolute value of the matrix element P in eq. (5) with the help of eq. (12). We obtain $P^2 = 0{\cdot}44$ atomic units.

The average hole mass m_{v1} is 0·18 m_0 as determined by thermo-electric power data[7] and data on the deionization of holes.[8] Cyclotron resonance[1] measurements also suggest a mass of this same value. The value 0·18 m_0 is not inconsistent with the binding energy of holes.[8] If we take the (110) direction as a good "average" direction we have from eq. (22) the relation

$$\frac{\hbar^2}{2}\left(\frac{1}{m_{v1}}+\frac{1}{m_0}\right)=\left|\frac{3}{8}B+\frac{1}{2}C+D\right|. \qquad (34)$$

The constant C may result from bands derived from atomic p functions, B requires d functions, D requires f functions. D is presumably small, but B and C may be comparable, according to the results in germanium.

No evidence relating to the size of the matrix elements F and G is available at present. Since F and G arise from the same band as C it is reasonable to suppose that they are comparable in magnitude. However, G would be zero by symmetry in a diamond structure. The constant G is important because it gives the splitting of the two conduction bands. A rough estimate of this effect gives a splitting of the order 0·1 eV at the Fermi level in n-type material with 10^{19} free electrons assuming $G \sim C$. A measurement of this splitting would aid in the determination of G.

Another important constant is the coefficient K of the linear k terms in the valence band structure. The constant K_2 appears to be considerably larger than K_1. An extremely rough estimate of the magnitude of K_2 was made which gave $K_2 \sim 2\times10^{-3}$ atomic units. This value corresponds to a maximum in the valence band ~10^{-4} eV above the energy at $k=0$ according to eq. (32). The value of k at the maximum is about 0·3 per cent of the value at the zone edge.

This estimate was made using published HARTREE[18] functions to calculate the spin-orbit matrix element for As and arbitrarily extrapolating to Sb by multiplying by the ratio of the atomic spin-orbit splittings of antimony and arsenic. The Γ_3 band involved in the estimate of K_2 was taken to be the 4d band of Sb. The 4d band was assumed to be about 30 eV below the valence band.* The matrix element of P between the valence band wave function and the 4d wave function was arbitrarily taken to be of the order $P \sim 0{\cdot}7$ atomic units. The spin-orbit interaction couples the 4d wave function to the 5d part of the valence band wave function. It was assumed that 50 per cent of the valence band wave function had 5d character.

At present there appears to be no reliable experimental evidence which bears on the magnitude of the linear k terms. Cyclotron resonance measurements should give the most reliable information.

* X-ray data shows the 4d–5s energy separation to be 26 eV in antimony.

At temperatures low compared to the energy of the valence band maximum, cyclotron resonance should show ellipsoids of revolution in the (111) direction. The ellipsoidal masses are given approximately by eq. (33). The longitudinal mass should be close to the average valence band mass determined from higher temperatures. The size of the transverse mass cannot be estimated accurately since the relative sizes of the constants B and C is not known. If C is large compared to B the ellipsoids will be prolate with an axis ratio of the order of 5. If B is large compared to C the extremum in the (111) direction will be a saddle point rather than a maximum.

At temperatures high compared to the energy of the valence band maximum the cyclotron resonance results should be similar to p-type germanium.

The cubic of eq. (10) has been solved numerically with the parameters $P^2 = 0{\cdot}44$ atomic units, $E_G = 0{\cdot}23$ eV, $\Delta = 0{\cdot}9$ eV. The heavy mass band E_{v1} was taken to have an average curvature corresponding to an effective mass $m_{v1} = 0{\cdot}18\ m_0$. The magnitude of the heavy mass m_{v1} gives evidence of significant perturbing effects from higher bands. These higher bands will have an important effect on the conduction band structure for energies of the order 0·5 eV above the bottom of the band. The bands resulting from the cubic equation (10) were corrected for higher bands according to eq. (19). The correction was made taking (110) as an average direction and assuming $C \gg B$, D. C can then be evaluated from $m_{v1} = 0{\cdot}18\ m_0$ using eq. (22). The assumption $C \gg B$ is very questionable. If $B \gg C$, the perturbation correction is one-third of the correction assuming $C \gg B$. Actually F and G are probably of the same order as C since they come from the same band but they are not included in these calculations for lack of information.

The computed band structure is shown in Fig. 1. The conduction band before correcting for the constant C is shown as a dashed curve. The nonparabolic nature of the conduction band for energies of the order 0·5 eV above $k = 0$ is quite apparent.

The nonparabolic region of the conduction band is associated with a wave function having a large admixture of p-symmetry. The coefficients a_i, b_i, c_i of eq. (14) as determined from the Hamiltonian of eq. (4) are shown in Fig. 2.

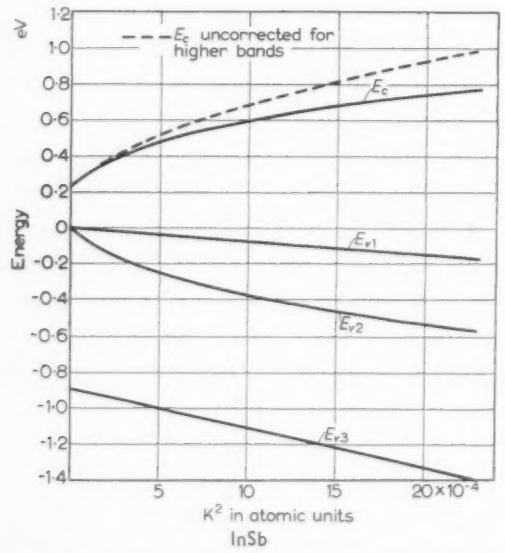


FIG. 1. Valence and conduction band energies vs.k^2 for an average direction in indium antimonide. The splitting of the double degeneracy of the bands is not given by these calculations.

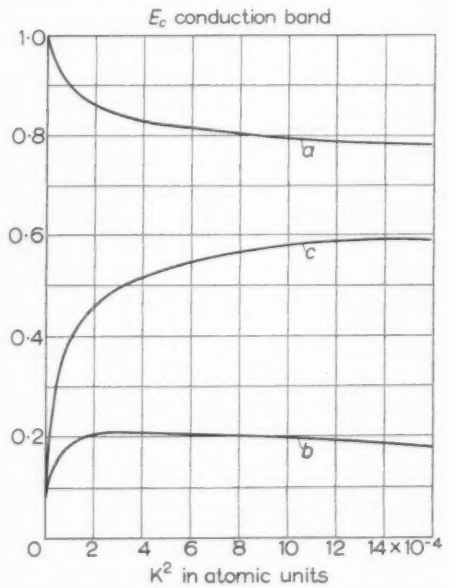


FIG. 2 (*a*).

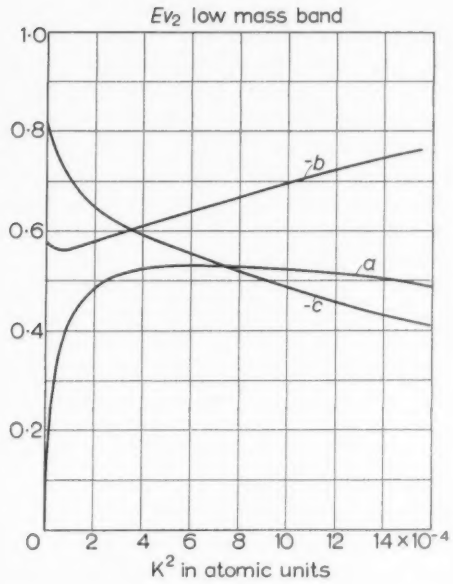


FIG. 2 (*b*).

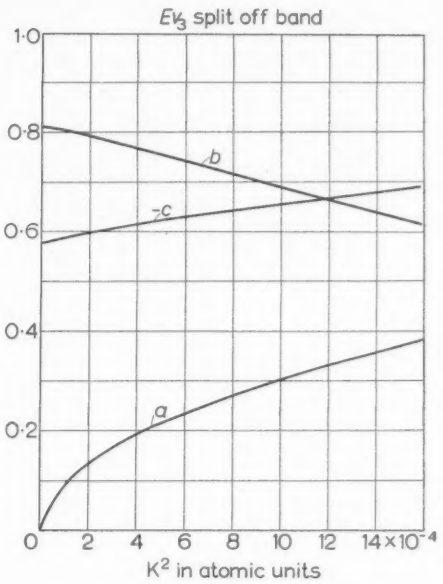


FIG. 2 (*c*).

FIG. 2. Wave function coefficients vs.k^2 in indium antimonide. See eq. (14) for definition of a, b, c.

4. OPTICAL ABSORPTION

Since we know the matrix element P of eq. (5) we can compute the fundamental optical absorption of InSb using the wave functions of eq. (14).

The absorption constant α may be written

$$\alpha = \frac{4\pi^2 e^2 \hbar}{m^2 cnE} \sum_j M_j^2 \rho_j(E) \tag{35}$$

where n is the index of refraction, E is the energy of the photon absorbed, M_j^2 is the square of the optical matrix element averaged over direction and summed over degenerate bands. The summation index j refers to the three twofold valence bands E_{v1}, E_{v2}, E_{v3}. The value of M_j^2 is computed from eqs. (14) and (5)

$$M_j^2 = \frac{2m^2P^2}{3\hbar^2}\{(a_c c_j + c_c a_j)^2 + (a_c b_j - b_c a_j)^2\}. \tag{36}$$

The coefficients a, b, c are defined by the eqs. (15). a_c, b_c, c_c refer to the conduction band. For the heavy mass band, E_{v1}; $a = 0$, $b = 1$, $c = 0$. Values of these coefficients are given in Fig. 2.

The density of states ρ_j is given by

$$\rho_j = k^2/2\pi^2|(dE_c/dk) - (dE_j/dk)|. \tag{37}$$

The optical absorption has been computed using the matrix element $P^2 = 0{\cdot}44$ atomic units determined from the cyclotron resonance mass and using the computed band structure shown in Fig. 1. The accuracy of the optical absorption calculations is determined by the accuracy of the band structure calculations. The results are shown in Fig. 3 together with the experimental data of FAN and GOBELI.[4] The theoretical calculations apply to $T = 0$°K whereas the experimental data were taken at room temperature. The entire theoretical curve was shifted by a constant energy 0·06 eV, corresponding to a room temperature band gap of 0·17 eV. Lattice dilation and phonon interaction effects also influence the temperature dependence. We have not included these effects; consequently the best test of the present theory will be given by low temperature optical absorption results when they become available. The theoretical curve agrees with the experimental curve except that it is about a factor of 1·5 too low in

absolute value. It is felt that the extent of the agreement between theory and experiment constitutes good evidence in support of the major assumptions about band structure made in this paper.

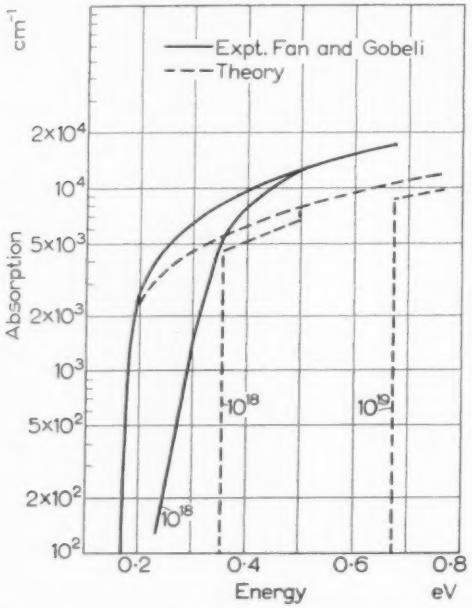


FIG. 3. Fundamental optical absorption in indium antimonide at room temperature. Concentration in impure specimens is noted on graph. Experimental data are due to FAN and GOBELI.[4]

The effect of n-type impurities on the optical absorption can be taken into account in the manner suggested by BURSTEIN.[19] In impure n-type specimens no bound states exist because the impurities are too close together. At absolute zero the electrons fill a sphere of radius k_F in k space according to the formula

$$n = k_F^3/3\pi^2. \tag{38}$$

In the approximation of direct transitions at $T = 0°K$ the electrons in the conduction band serve to cut off the absorption resulting from k values less than k_F. This effect is shown in Fig. 3 for concentrations of 10^{18} and 10^{19}. The main edge results from cutting off the absorption due to the heavy-mass band. The small additional absorption due to the light-mass band is cut off at higher energy. The absorption has been shifted to account for the change in band gap with temperature. In this case comparison is further complicated by the effect of temperature in allowing direct transitions to states below the Fermi level as discussed by KAISER and FAN.[20] The agreement of theory and experiment on the location of the cut off appears to be fairly good. It would be very desirable to be able to make comparison to data taken at lower temperatures.

The effect of phonons and impurity scattering is to allow indirect transitions which will round off the corners shown in Fig. 3. The threshold for indirect transitions will be at an energy lower than the threshold for direct transitions by the amount

$$\Delta E = \hbar^2 k_F^2/2(0{\cdot}18m_0). \tag{39}$$

No account has been taken of the effect of impurities on the band structure. If this effect is calculated using perturbation theory by the method of PARMENTER,[12] the change in band structure is found to be very small.

Acknowledgements—The author is indebted to Professor H. Y. FAN and G. W. GOBELI for the use of their experimental data. He is also very grateful to H. EHRENREICH for many valuable discussions and to F. HAM for useful comments. He wishes to thank Miss ANN WARNER for performing the numerical calculations.

REFERENCES

1. DRESSELHAUS G., KIP A. F., KITTEL C., and WAGONER *Phys. Rev.* **98,** 556 (1955).
2. TANENBAUM M., PEARSON G. L., and FELDMAN W. L. *Phys. Rev.* **93,** 912 (1954).
3. FREDERIKSE H. P. R. and HOSLER W. R. *Bull. Amer. Phys. Soc.* **6,** 298 (June 1956).
4. FAN H. Y. and GOBELI G. W. *Bull. Amer. Phys. Soc.* **3,** 111 (March 1956).
5. ROBERTS V. and QUARRINGTON J. E. *J. Electronics* **1,** 152 (1955).
6. DRESSELHAUS G. *Phys. Rev.* **100,** 580 (1955). PARMENTER R. H. *Phys. Rev.* **100,** 573 (1955).
7. WEISS H. *Z. Naturforschung* **11a,** 131 (1956).
8. HROSTOWSKI H. J., MORIN F. J., GEBALLE T. H., and WHEATLEY G. H. *Phys. Rev.* **100,** 1672 (1955).
9. WILLARDSON R. K., HARMAN T. C., and BEER A. C. *Phys. Rev.* **96,** 1512 (1954).
10. MACFARLANE G. C. and ROBERTS V. *Phys. Rev.* **97,**

VOL. 1 1956/57

1714 (1955). BLOUNT E., CALLAWAY J., COHEN M., DUMKE W., and PHILLIPS J. *Phys. Rev.* **101,** 563 (1956). POTTER R. F. *Phys. Rev.* (to be published).
11. STERN F. and TALLEY R. M. *Phys. Rev.* **100,** 1638 (1955)
12. PARMENTER R. H. *Phys. Rev.* **97,** 587 (1955).
13. KANE E. O. *J. Phys. Chem. Solids* **1/2,** 82 (1956). This paper contains a bibliography of work on the $k \cdot p$ approach.
14. DRESSELHAUS G., KIP A. F., and KITTEL C. *Phys. Rev.* **98,** 368 (1955).
15. ELLIOTT R. J. *Phys. Rev.* **96,** 266 (1954).
16. KAHN A. H. *Phys. Rev.* **97,** 1647 (1955).
17. SPITZER W. G. and FAN H. Y. *Phys. Rev.* **99,** 1893 (1955).
18. HARTREE W., HARTREE D. R., and MANNING M. F. *Phys. Rev.* **59,** 299 (1941).
19. BURSTEIN E. *Phys. Rev.* **93,** 632 (1954).
20. KAISER W. and FAN H. Y. *Phys. Rev.* **98,** 966 (1955).

J. Phys. Chem. Solids. Pergamon Press 1957. Vol. 1. pp. 262–269.

SCATTERING OF CARRIERS BY IONIZED IMPURITIES IN SEMICONDUCTORS*†

F. J. BLATT‡

Physics Department, University of Illinois, Urbana, Illinois

(*Received* 30 *August* 1956)

Abstract—The cross-sections for scattering of carriers by ionized impurities have been calculated by the partial wave method. The potential used to describe the scattering center was the screened Coulomb potential suggested by BROOKS, HERRING, and DINGLE. The surfaces of constant energy in **k** space were assumed to be spheres. The results have been compared to those obtained by application of the Born approximation.

At small carrier energies, that is low temperatures, the Born approximation leads to incorrect results, the correct cross-sections being significantly less than the approximate results. In that energy range majority impurities scatter much more effectively than minority impurities. Under certain conditions transmission resonances can occur. It is unlikely, however, that these resonances could produce a pronounced effect on the mobility although they might be observable. Detailed calculations of the Hall and drift mobilities are presented for two special cases. It is found that the mobility exhibits a minimum at a rather low temperature. This minimum is probably obscured in practice by such effects as neutral impurity scattering and impurity band conduction. At higher temperatures the mobility increases with temperature less rapidly than $T^{3/2}$. In fact, the temperature dependence of the mobility cannot be expressed in terms of a simple power law although at high temperature it appears to approach a $T^{3/2}$ behavior. At temperatures between liquid air and room temperature the Hall and drift mobilities increase with temperature less rapidly than $T^{3/2}$. Significant deviations from the $T^{3/2}$ law occur even in that range in which the Born approximation gives the correct scattering cross-sections. The usual assumption of a $T^{3/2}$ law for ionized impurity scattering is believed to be due to unwarranted approximate procedures in the evaluation of certain integrals over the Boltzmann distribution.

VOL.
1
1956/57

INTRODUCTION

THE mobility of charge carriers in semiconductors is limited by various scattering mechanisms.[1, 2] At high temperatures scattering by lattice vibrations (phonons) is predominant. As the temperature is lowered lattice scattering decreases, the relaxation time going as $1/T\epsilon^{1/2}$, where ϵ is the kinetic energy of the carrier. At some temperature which, for a given substance depends on the concentration of impurities, scattering by ionized donors and acceptors becomes the paramount scattering mechanism. In germanium and silicon for reasonable impurity concentrations ionized impurity scattering is dominant between about hydrogen temperature and 100° to 200°K. At very low temperatures, the exact temperature region again depending on the impurity concentration and the extent of compensation, scattering by neutral impurities is primarily responsible for a finite relaxation time. Under certain circumstances other static imperfections such as dislocations may also limit the mobility.[3]

Lattice scattering has been discussed by SEITZ,[4] BARDEEN and SHOCKLEY,[5] DUMKE,[6] EHRENREICH and OVERHAUSER,[7] and others.[1] Scattering by neutral impurities was considered by ERGINSOY.[8] In this article we focus attention on scattering of charge carriers by ionized impurities.

Ionized impurity scattering in semiconductors was discussed first by CONWELL and WEISSKOPF.[9] In recent years their calculation has been reconsidered

* Supported by the Office of Naval Research.

† A report of the preliminary results of this investigation was presented at the American Physical Society Meeting, New York, January 1956; *Bull. Amer. Phys. Soc.* II, **1**, 48 (1956).

‡ Present address: Physics Department, Michigan State University, East Lansing, Michigan.

and improved by BROOKS,[10] HERRING,[11] and DINGLE.[12] In these treatments the surfaces of constant energy in $\boldsymbol{k}$ space were assumed to be spherical. HAM[13] has considered impurity scattering when the energy surfaces are ellipsoids and found that the relaxation time exhibits a pronounced anisotropy. The scattering potentials employed by these workers were different, though based in each case on the hydrogenic model. In all of these calculations, however, it was assumed that the Born approximation could be used to determine the total cross-section for scattering of carriers by the ionized impurities. We shall concern ourselves in this paper only with the question of the validity of the Born approximation when used in calculations of this sort.*

As is well known, the Born approximation is valid if the kinetic energy of the carriers is large compared to the "effective strength" of the scattering potential.[14] In our case this effective strength is of the order of the ionization energy of the impurity. It is obvious that as the temperature is lowered, reducing the average kinetic energy of the charge carriers, an increasingly large fraction of the carrier distribution will fail to satisfy the validity condition for the Born approximation. We must anticipate, therefore, that results derived by the use of this approximation will fail at sufficiently low temperatures.

To ascertain the error associated with the use of the Born approximation we have calculated the scattering cross-sections, weighed by the factor $(1-\cos\theta)$, for positively and negatively charged impurities employing the partial wave method. The results are compared to those found in the Born approximation.

Concurrently and independently SCLAR at the Naval Research Laboratory has considered the same problem.[15] SCLAR employed square well (or barrier) potentials and potentials of the form

$$V(r)=\frac{1}{r}(1-r/a),\ r\leqslant a;\quad V(r)=0,\ r\geqslant a.$$

These potentials have the advantage that the phase shifts can be determined analytically. We have used the screened Coulomb potential

$$V(r)=\frac{1}{r}\exp(-r/R) \tag{1}$$

following the work of BROOKS, HERRING, and DINGLE. Our results appear to be in qualitative agreement with those of SCLAR.

In the calculations which we report here we have retained the simplifying assumption that the energy surfaces are spherical. Work now in progress is aimed at removing this restriction in the framework of the partial wave method.

RESULTS AND DISCUSSION

The method of calculating the differential scattering cross-section by a partial wave analysis is discussed in standard references.[14, 16] HUANG[17] has shown that the total cross-section for momentum transfer

$$A=\int(1-\cos\theta)\sigma(\theta)\,d\omega \tag{2}$$

is readily obtained once the phase shifts are known.

When the scattering potential is of the form of Eq. (1) the calculation of the phase shifts is not feasible by analytical methods but requires numerical computation. The numerical work was performed with the aid of the ILLIAC, employing a suitable program developed and discussed by RUBENSTEIN, HUSE, and MACHLUP.[18]

The cross-section for the screened Coulomb potential in the Born approximation is given by[12, 19]

$$A=\frac{\pi}{2}\frac{Z^2e^4}{\kappa^2\epsilon^2}\left\{\ln(1+b)-\frac{b}{1+b}\right\}=\frac{\pi}{2}\frac{Z^2e^4}{\kappa^2\epsilon^2}g(b) \tag{3}$$

where $b=32\pi^2m^*R^2\epsilon/h^2$.

The screening radius R in Eq. (1) is a function of temperature and carrier concentration.[12] In the region of interest in which the electron distribution is nondegenerate the screening radius is given by

$$R\times 13{\cdot}1\times 10^8(\kappa T/n)^{\frac{1}{2}}\text{ atomic units.} \tag{4}$$

It is evident from inspection of the radial wave equation that by suitable scaling of the scattering

* These same considerations apply also in the case of electron-electron and electron-hole scattering. Here, the strength of the scattering potential as well as the energies involved are essentially the same as in scattering by singly charged ionized impurities, and, hence, the Born approximation is likely to be unreliable at low carrier energies.

potential the results can be used for arbitrary choices of the parameters κ (dielectric constant), m^* (effective mass), and Z (charge of the impurity center). The scaling factor we have chosen is

$$Q = 2{\cdot}8(\kappa/Zm^*). \tag{5}$$

The electron wave vector k and the screening radius R in atomic units are related to k' and R' by

$$k = k'/Q, \quad R = R'Q. \tag{5a}$$

The scattering cross-section A in atomic units is given by

$$A = A'Q^2/0{\cdot}28 = 28(\kappa/Zm^*)^2A' \tag{5b}$$

The results are shown in Figs. 1 to 6. The curves labeled (+), (−) and B correspond to the partial wave results for attractive and repulsive potentials and to the results obtained by use of the Born approximation, respectively. In the Born approximation attractive and repulsive potentials lead to identical cross-sections.

The partial wave results are quite accurate for $k' \leqslant 1$. As the electron energy is increased, more and more partial waves contribute to the scattering cross-section. The ILLIAC program could not integrate the radial wave equation for angular momenta in excess of $l = 5$. As phase shifts for $l > 5$ became significant they were calculated in the Born approximation[16] as long as that procedure appeared consistent (generally to $k'^2 \sim 3$). Thereafter, the curves for A' were extrapolated. It is unlikely that the results for $k'^2 > 3$ are in error by more than about 25 per cent.

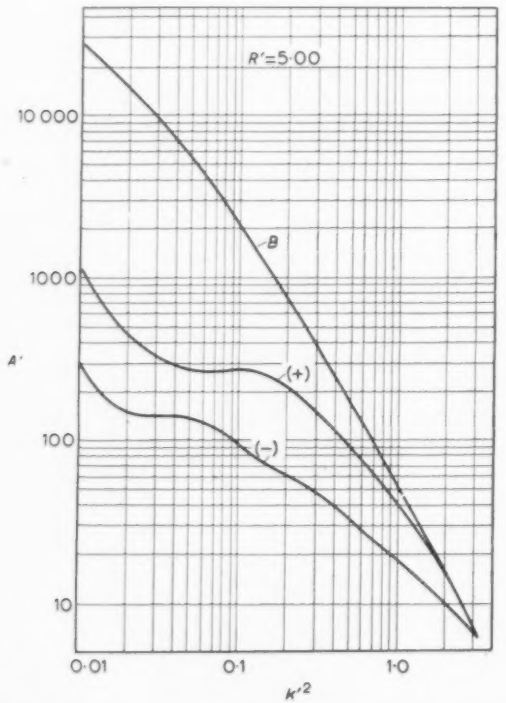


FIG. 1. The normalized total scattering cross-section as a function of the normalized carrier energy for $R' = 5{\cdot}0$. The curve labeled (+) shows the cross-section for an attractive potential, that labeled (−) shows the cross-section for a repulsive potential. The curve labeled **B** shows the cross-section obtained by application of the Born approximation.

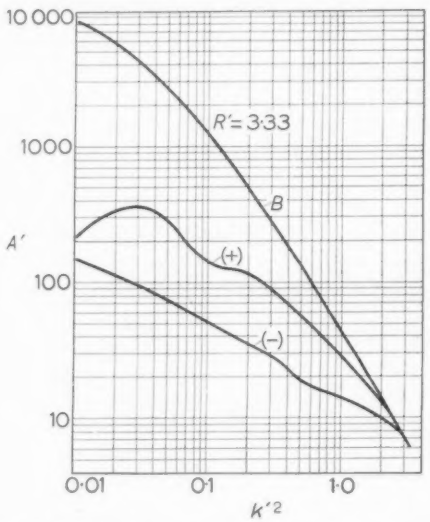


FIG. 2. The normalized total scattering cross-section as a function of the normalized carrier energy for $R' = 3{\cdot}33$. The curve labeled (+) shows the cross-section for an attractive potential, that labeled (−) shows the cross-section for a repulsive potential. The curve labeled **B** shows the cross-section obtained by application of the Born approximation.

We have made no attempt to present a similar comparison with the Conwell-Weisskopf results, even though their work has been used quite extensively until recently. The difficulty is that no unique potential equivalent to the Conwell-Weisskopf method can be obtained. These workers

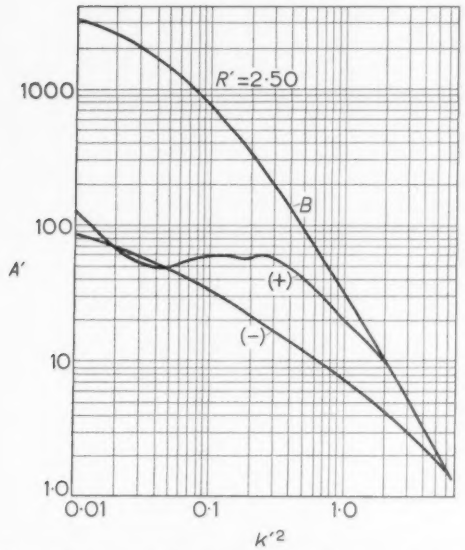


FIG. 3. The normalized total scattering cross-section as a function of the normalized carrier energy for $R' = 2{\cdot}5$. The curve labeled (+) shows the cross-section for an attractive potential, that labeled (−) shows the cross-section for a repulsive potential. The curve labeled **B** shows the cross-section obtained by application of the Born approximation.

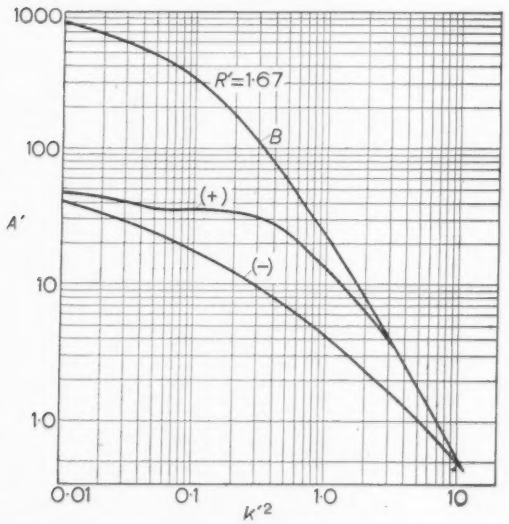


FIG. 4. The normalized total scattering cross-section as a function of the normalized carrier energy for $R' = 1{\cdot}67$. The curve labeled (+) shows the cross-section for an attractive potential, that labeled (−) shows the cross-section for a repulsive potential. The curve labeled **B** shows the cross-section obtained by application of the Born approximation.

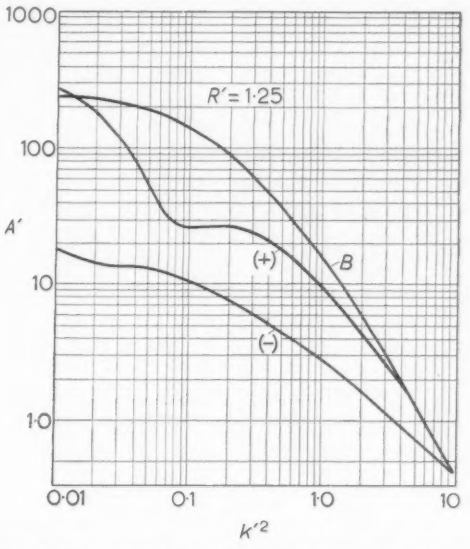


FIG. 5. The normalized total scattering cross-section as a function of the normalized carrier energy for $R' = 1{\cdot}25$. The curve labeled (+) shows the cross-section for an attractive potential, that labeled (−) shows the cross-section for a repulsive potential. The curve labeled **B** shows the cross-section obtained by application of the Born approximation.

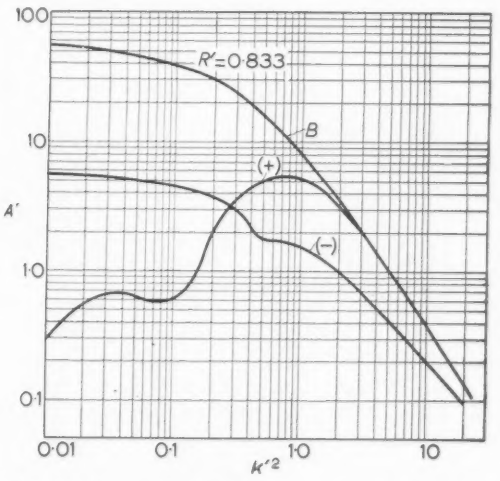


FIG. 6. The normalized total scattering cross-section as a function of the normalized carrier energy for $R' = 0{\cdot}833$. The curve labeled (+) shows the cross-section for an attractive potential, that labeled (−) shows the cross-section for a repulsive potential. The curve labeled **B** shows the cross-section obtained by application of the Born approximation.

used a pure Coulomb potential from the origin to infinity but avoided the well-known divergence difficulty by cutting off the impact parameter when integrating the differential scattering cross-section over solid angle. In other words some electrons were assumed to feel the pure Coulomb potential out to infinite distance, while others, beyond the specified impact parameter, were assumed to be outside the range of the potential.*

There are a number of features of interest. First, as we anticipated, the divergence between the partial wave and the Born approximation cross-sections is most pronounced for carriers of small energy (small values of k'^2). In that region the Born approximation consistently overestimates the true cross-section. At high energies the results for the two methods appear to converge.

Second, a repulsive potential generally leads to a smaller cross-section than an attractive potential. This means that, in the range where the Born approximation fails, ionized minority impurities are considerably less effective as scattering centers than are ionized majority impurities. The reduction in mobility due to partial compensation is, therefore, not so severe as one might expect on the basis of the Born approximation, in which both types of impurities scatter with equal strength.

Third, transmission resonances occur under certain circumstances (see Fig. 6). These resonances, corresponding to the well-known Ramsauer effect in scattering of electrons by gaseous atoms, depend fairly sensitively on the value of the screening radius. One might think, at first, that such resonances should be readily observable leading to abnormally high mobilities. To be sure the mobility will be enhanced by these resonances, but the following consideration should be kept in mind.

Impurities in any real specimen are distributed at random, not uniformly. Therefore, even with given impurity and carrier concentrations, the screening radius for various impurities in the sample will fluctuate about an appropriate mean. If the Born approximation is used to calculate the relaxation time it is possible to devise suitable methods which take account of such a random distribution.[10] We have considered ways of performing analogous calculations in the framework of the partial wave method but have not found a solution to this problem. Although we cannot, therefore, provide a satisfactory answer to the question concerning the influence of randomness of the impurity centers on the mobility, it seems to us that though resonances may make some ionized impurities ineffective scattering centres, such a situation cannot prevail throughout the sample, other impurities scattering more effectively.

Similar resonances have also been found by SCLAR, who would attach considerably more importance to them than we do. He has suggested,[20] for example, that transmission resonances may account for the effects at low temperatures which are generally explained in terms of an impurity band model.[21] Quite apart from the point raised in the previous paragraph, it seems difficult to account in this manner for the observed increase of impurity band effects with ever increasing impurity concentration. The resonances are sensitive to the screening radius and hence to impurity concentration and should, therefore, result in a behavior characteristic of a resonance effect.

The calculation of Hall and drift mobilities and of the coefficient of magnetoresistance and their temperature dependences in the ionized impurity scattering regime requires integrations over the Boltzmann distribution. The quantities that are needed are averages of τ, τ^2 and τ^3, where τ is the relaxation time.[1] In terms of the cross-sections we would have to calculate integrals of the form

$$I_1 = \int_0^\infty A^{-1}(x) x e^{-x}\, dx,$$

$$I_2 = \int_0^\infty A^{-2}(x) x^4 e^{-x}\, dx,$$

$$I_3 = \int_0^\infty A^{-3}(x) e^{-x}\, dx,$$

in which $x = \epsilon/k_0 T$. Since a new set of integrations has to be performed for every choice of impurity concentrations the task of preparing a fairly complete set of mobility curves is considerable. For any specific set of conditions the curves of Figs. 1 to 6 should permit one to calculate the mobility. In order to exhibit the changes which

* I am grateful to Dr. D. L. DEXTER for pointing this out to me.

VOL. 1 1956/5

arise when the partial wave results are employed we present curves for Hall and drift mobilities calculated for the following conditions: $m^* = 0{\cdot}33$, $\kappa = 11{\cdot}8$, $n = 10^{16}$ and $n = 2\times10^{16}$. The number of acceptors was assumed to be zero. This is, of course, a highly artificial assumption. However, as we have pointed out earlier, minority impurities are significantly less effective than majority impurities in the range where the Born approximation fails. Nevertheless, for any real case, compensation must be taken into account. The values for m^* and κ are probably appropriate for n-type silicon. We do not show curves for the magnetoresistive coefficient because the magnetoresistance is very sensitive to the shape of the energy surfaces.[22] The assumption of spherical energy surfaces which we have employed must make such results of questionable value.

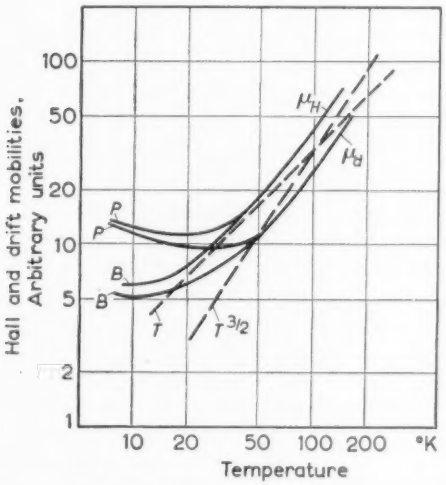


FIG. 7. Hall and drift mobilities as functions of temperature for the case $m^* = 0{\cdot}33$, $K = 11{\cdot}8$, $n = 10^{16}$, and no compensation. The curves labeled **P** were obtained using the results of the partial wave method. The curves labeled **B** were obtained using the results of the Born approximation. Also shown are the two straight lines corresponding to a $T^{3/2}$ and a linear temperature dependence.

In Figs. 7 and 8 are shown the Hall and drift mobilities for the two conditions specified above. The curves labeled ***B*** were obtained by use of the Born approximation cross-sections. The curves marked ***P*** show the mobilities calculated from the partial wave cross-sections.

In order to avoid extraneous discrepancies we have not resorted to the usual approximation[1] of neglecting the energy variation of the term g(b) in the Born approximation cross-section, eq. (3).

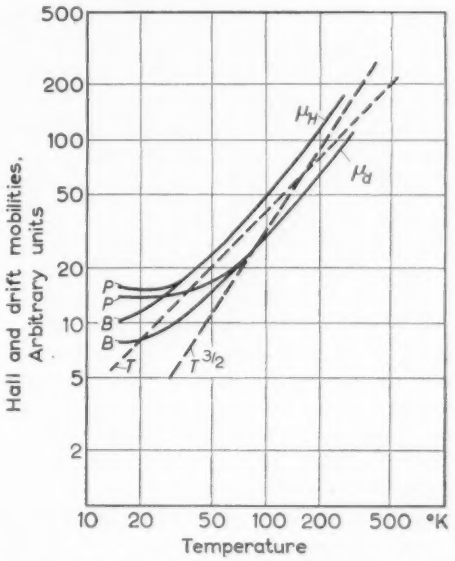


FIG. 8. Hall and drift mobilities as functions of temperature for the case $m^* = 0{\cdot}33$, $K = 11{\cdot}8$, $n = 2\times10^{16}$, and no compensation. The curves labeled **P** were obtained using the results of the partial wave method. The curves labeled **B** were obtained using the results of the Born approximation. Also shown are the two straight lines corresponding to a $T^{3/2}$ and a linear temperature dependence.

This term is generally removed from under the integral sign in order to facilitate integration. In the case of the curves presented in Figs. 7 and 8 the integrals were obtained with the aid of a planimeter.

Several features are of interest here. First, the mobilities in that region of temperature in which the Born approximation fails are larger than predicted by that approximation. This is an obvious consequence of the smaller cross-sections in the partial wave method.

Second, at low temperatures the mobility levels off and at very low temperatures increases as the temperature decreases. This effect is more pronounced when the partial wave cross-sections are used although it is also apparent in the Born approximation. The reason for this change in the

temperature dependence is that the scattering cross-sections become independent of energy at low energies in the Born approximation, as may be verified by expanding $g(b)$ for small values of b. This same behavior also is approximately obeyed (except when resonances occur) if the partial wave method is employed, although in that case the cross-sections appear to level off already at somewhat higher energies.

Third, even in the region of higher temperatures in which both methods of calculating the cross-sections give identical results, the mobilities do not follow the predicted $T^{3/2}$ law.

Indeed, the Hall and drift mobilities appear to follow no simple power law dependence on temperature. The slopes of the tangents to the curves of Figs. 7 and 8 increase monotonically as the temperature increases and tend to approach the value 3/2 at high temperature. Near liquid-air temperature the slopes of the tangents are significantly less than 3/2. It would appear that the deviation from the $T^{3/2}$ dependence arises because we have performed the integrations over the Boltzmann distribution without first removing the term $g(b)$ from the integrands.

DEBYE and CONWELL[23] concluded on the basis of their experimental results that the temperature dependence of the mobility in the ionized impurity scattering regime is indeed less rapid that $T^{3/2}$, and "that the exponent of T should be closer to 1·0 than 1·5". Our results are consistent with these observations.

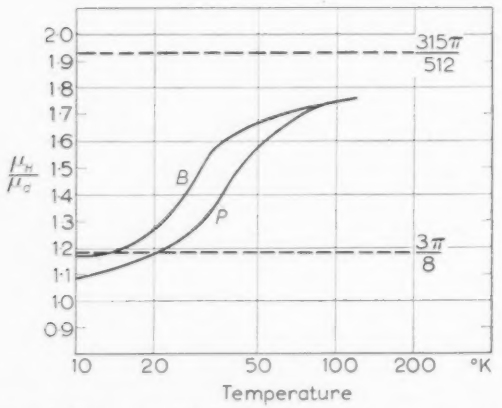


FIG. 9. The ratio of Hall-to-drift mobilities as a function of temperature for the case $m^* = 0{\cdot}33$, $K = 11{\cdot}8$, $n = 10^{16}$, and no compensation. The curve labeled **P** corresponds to the partial wave cross-sections. The curve labeled **B** corresponds to the Born approximation cross-sections. Also shown are the values $315\pi/512$ and $3\pi/8$.

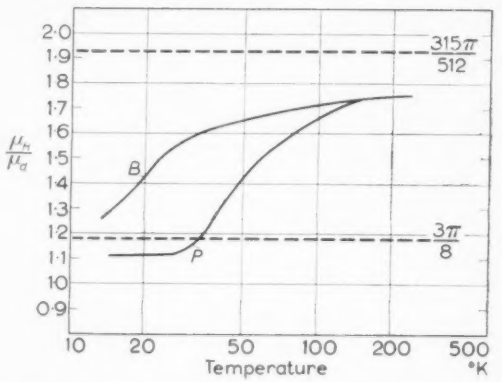


FIG. 10. The ratio of Hall-to-drift mobilities as a function of temperature for the case $m^* = 0{\cdot}33$, $K = 11{\cdot}8$, $n = 2\times10^{16}$, and no compensation. The curve labeled **P** corresponds to the partial wave cross-sections. The curve labeled **B** corresponds to the Born approximation cross-sections. Also shown are the values $315\pi/512$ and $3\pi/8$.

Figs. 9 and 10 show the temperature dependence of the ratio of Hall-to-drift mobilities. The two limiting values $315\pi/512$ and $3\pi/8$ are also indicated. The lower limit is approached at low temperatures where, as we remarked earlier, the cross-sections are independent of electron energy. The fact that at high temperatures the mobility ratio only slowly approaches but does not attain the value $315\pi/512$ is also consistent with a temperature dependence of the mobility slower than $T^{3/2}$.

CONCLUSION

We have presented here in graphical form results of calculations of the cross-sections for scattering of carriers by ionized impurities in semiconductors. The calculations were motivated by the belief that in certain temperature ranges the Born approximation is likely to fail, and that the result derived by the application of this approximation should be revised. We have used the partial wave method and have compared our cross-sections with those obtained in the Born approximation. The following conclusions may be drawn.

VOL. 1 1956/57

1. The Born approximation is good only in the higher temperature region but fails at low temperatures. When this approximation loses validity it considerably overestimates the scattering cross-sections.

2. Attractive centers scatter more effectively than repulsive centers. Therefore, minority impurities are not as effective in limiting the mobility at low temperatures as are majority centers.

3. Under certain conditions transmission resonances may occur. Although such resonances may give rise to an enhancement of the mobility it is unlikely that they could have a very profound effect because the random distribution of impurity centres in a real specimen will probably smear out the resonance phenomenon.

4. In the ionized impurity scattering regime the mobility should attain a minimum at a rather low temperature, although neutral impurity scattering and impurity-band conduction may make this effect unobservable.

5. At higher temperatures the temperature dependence of the Hall and drift mobilities is less rapid than $T^{3/2}$. Instead, the mobilities can not be expressed in the form $\mu = \mu_0 T^n$ although at high temperatures this relation, with $n = 3/2$, appears to be approached. At lower temperatures the temperature dependence of the mobilities is less rapid than $T^{3/2}$, a result which is consonant with experimental observation. This "slower-than-predicted" temperature dependence occurs also if the Born approximation is used. It is, therefore, not due to a failure of that approximation but rather the result of retaining the term $g(b)$ of eq. (3) in the appropriate integrands. In other words, previous expressions for the mobility in the ionized impurity scattering regime have not given the true temperature variation because factors which depend on temperature and carrier energy were not considered in sufficient detail.

6. The ratio of Hall-to-drift mobilities in the temperature range where ionized impurity scattering is dominant is not independent of temperature. This ratio actually decreases from a value somewhat less than the "theoretical" value of $315\pi/512$ to the value $3\pi/8$ as the temperature is reduced.

Acknowledgements—The author is grateful to Dr. S. MACHLUP and to the staff of the University of Illinois Computer Laboratory for their help in the numerical computations. Discussions with Dr. N. SCLAR about his work are also appreciated. It is a pleasure to thank Dr. H. BROOKS for valuable comments on the original manuscript.

REFERENCES

1. BROOKS H. Theory of the electrical properties of germanium and silicon in *Advances in Electronics and Electron Physics*, Vol. 7. Academic Press, Inc., New York (1956).
2. SHOCKLEY W. *Electrons and Holes in Semiconductors* D. Van Nostrand Co., New York (1950).
3. READ W. T. *Phil. Mag.* **45**, 119, 775 (1954); **46**, 111 (1955).
4. SEITZ F. *Phys. Rev.* **73**, 549 (1948).
5. BARDEEN J. and SHOCKLEY W. *Phys. Rev.* **80**, 72 (1950)
6. DUMKE W. P. *Phys. Rev.* **101**, 531 (1956).
7. EHRENLEICH H. and OVERHAUSER A. W. to be published.
8. ERGINSOY C. *Phys. Rev.* **79**, 1013 (1950).
9. CONWELL E. M and WEISSKOPF V. F. *Phys. Rev.* **77**, 388 (1950).
10. BROOKS H. *Phys. Rev.* **83**, 879 (1951). See also ref. 1.
11. HERRING C. unpublished results.
12. DINGLE R. B. *Phil. Mag.* **46**, 831 (1955).
13. HAM F. S. *Phys. Rev.* **100**, 1251 (1955). See also ref. 1, page 158.
14. MOTT N. F and MASSEY H. S. W. *The Theory of Atomic Collisions*, 2nd ed. (Oxford University Press, London, 1949).
15. SCLAR N. *Bull. Amer. Phys. Soc.* II, **1**, 48 (1956).
16. SCHIFF L. I. *Quantum Mechanics* (McGraw-Hill Book Co., New York, 1949).
17. HAUNG K. *Proc. Phys. Soc.* (London) **60**, 161 (1948).
18. RUBENSTEIN R., HUSE M., and MACHLUP S. *Math. Tables and Other Aids to Computation* **10**, 30 (1956).
19. MOTT N. F. *Proc. Cambridge Phil. Soc.* **32**, 281 (1936).
20. SCLAR N. (to be published); also private communication.
21. FRITZSCHE H. and LARK-HOROVITZ K. *Physica* **20**, 834 (1954).
22. ABELES B and MEIBOOM S. *Phys. Rev.* **95**, 31 (1954). SHIBUYA M. *Phys. Rev.* **95**, 1385 (1954).
23. DEBYE P. P. and CONWELL E. M. *Phys. Rev.* **93**, 693 (1954).

J. Phys. Chem. Solids. Pergamon Press 1957. Vol. 1. pp. 270–274.

SIGN REVERSAL OF THE HALL EFFECT IN RHODIUM

B. R. COLES*

Carnegie Institute of Technology, and Ford Motor Company, Scientific Laboratories, Dearborn Michigan, U.S.A.

AND J. C. TAYLOR†

Physics Department, Carnegie Institute of Technology, Pittsburgh, Pennsylvania, U.S.A.

(*Received* 20 *September* 1956)

VOL. 1 1956/57

Abstract—Measurements of the Hall effect in rhodium have been made at temperatures between 4·2°K and 300°K in fields up to 22 kilogauss. At room temperature the Hall voltage is linear in field, but at lower temperatures non-linearity becomes apparent at higher fields. At 20·4° and 4·2°K the Hall voltage has a maximum at about 8 kilogauss, decreasing with further increase of field and changing sign at about 18 kilogauss. The low-field Hall coefficient (which has the value $+5{\cdot}02\times10^{-13}$ volt cm amp^{-1} gauss^{-1} at 300°K) varies only slightly with temperature, having a shallow minimum at about 60°K. The results can be fitted quite closely by expressions of the form given by theoretical models, and the significance of the values required for the parameters in these expressions is discussed in terms of reasonable models for the electronic structure of rhodium. The results of magneto-resistance measurements on the same specimen are not as conveniently described by theoretical expressions as are the Hall effect data.

1. INTRODUCTION

MANY of the physical properties of the first transition group metals and their alloys have been investigated in detail, but fewer data are available for metals of the later transition groups. The results of thermo-electric measurements made on alloys of palladium with silver and with rhodium have been described elsewhere[1] and the present paper presents some results obtained in the measurement of the Hall effect in such materials.

In a recent article[2] attention has been drawn to important differences between the first and later transition groups, and further measurements (especially of those properties which can yield information about the electronic structures of these metals and their alloys) seem called for. Hall effect measurements cannot always be given a simple interpretation in terms of the electronic structure of a metal, and in alloys the situation appears even more uncertain.[3] Considerable information is now available, however, on the ordinary Hall effects of metals and alloys of the first transition group[4, 5] and it seems desirable to compare with such data the results of measurements made on second transition group materials. There is a great deal of evidence to suggest that nickel and palladium, the end members of the first and second transition groups, have very similar electronic structures in the metallic state, and the similarity of their Hall coefficients in both sign and magnitude is understandable. The results presented below suggest strongly that this similarity does not necessarily hold for earlier members of the two groups.

2. MATERIAL AND METHODS

The rhodium used in this investigation was kindly loaned to the authors by the Mond Nickel Company (Precious Metals Division) and was of a purity exceeding 99·9 per cent; the dimensions of the specimens were approximately 7 cm $\times$ 1cm $\times$ 0·1 cm. The measurements were made by a standard type of d.c. method, the Hall potential being measured by a Rubicon microvolt potentiometer in conjunction with a twin photo-cell galvanometer amplifier. A double probe system

* Now at Imperial College, London S.W.7.

† Now at AEI Research Laboratory, Aldermaston, Berkshire.

was used for the potential leads and a small potential was present in zero magnetic field. Measurements were therefore made by reversing the field direction and noting the change in the potential across the Hall probes. These results were checked by reversing the specimen current and correcting the potential changes for the zero field change, this correction being quite small at the lowest temperatures. For all measurements in liquid gas baths thermo-electric and thermo-magnetic contributions to the measured potential constituted less than 1 per cent of the total. The specimen holder and Dewar assembly was modelled upon that used by Dr. P. COHEN for measurements on copper–nickel alloys at low temperatures (unpublished Ph.D. thesis, Carnegie Institute of Technology, 1955) and the authors acknowledge with gratitude the assistance of Dr. F. E. ALLISON in the design and construction of the apparatus. The magnet used was an A. D. Little electromagnet with tapered iron–cobalt pole pieces, and the specimen currents used were of the order of 5 ampères.

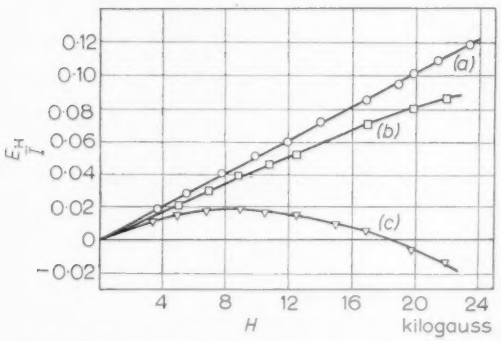


FIG. 1. Hall potential of rhodium per unit specimen current in microvolts per ampère as a function of magnetic field in kilogauss; (a) at 300°K (b) at 77°K (c) at 20·4°K. Specimen thickness 0·1 cm.

3. RESULTS

At room temperature the Hall coefficient (R) was found to be independent of the field, but at temperatures of and below that of liquid nitrogen a field dependence was observed. In Figure 1 the Hall potential for unit specimen current is plotted as a function of the field for three different temperatures; and in Figure 2 the Hall coefficient, together with theoretical curves for this quantity,

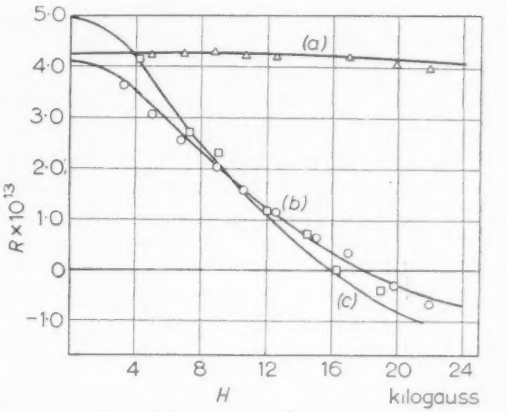


FIG. 2. Hall coefficient (R) of rhodium in volt cm amp^{-1} gauss^{-1} as a function of magnetic field; (a) at 77°K (b) at 20·4°K (c) at 4·2°K. The solid curves are those given by the two-band model using the values of the parameters given in the text.

Table 1. *Hall Coefficient* (R) *of rhodium at various temperatures in volt cm amp*$^{-1}$ *gauss*$^{-1}$ R_0 *at zero magnetic field,* R_{10} *at* 10 *kilogauss,* R_{20} *at* 20 *kilogauss*

T(°K)	300	77	65	20·4	17	4·2
$R_0\times10^{13}$	5·02	4·24	3·9	4·1	4·2	5·0
$R_{10}\times10^{13}$	5·02	4·20	3·87	1·70	1·78	1·88
$R_{20}\times10^{13}$	5·02	4·02	3·30	−0·37	−0·38	−0·56
$r_T\times10^3$	1134·0	113·7	68·3	10·38	9·66	9·32

is plotted against the field. Table 1 shows the values of the Hall coefficient for fields of 0, 10, and 20 kilogauss for all the temperatures employed, these values being taken from appropriate curves through the experimental points of plots like Figure 2. The last row of this table gives the ratio (r_T) of the electrical resistivity at the temperature T to that at 273°K; the value of this quantity at helium temperatures is, of course, a measure of the purity of the specimen. The values for the zero field Hall coefficient R_0 are somewhat uncertain at the lowest temperatures since they involve extrapolations, and those quoted are the results of extrapolations of the form suggested by the theoretical curves which best fit the experimental points. The room temperature value of R is in excellent agreement with that given by GEHLOFF and JUSTI.[6]

When the Hall coefficient was found to show such a strong field-dependence extra probes were attached to one side of the specimen about 5 cm apart, and the change of resistance with transverse magnetic field was measured. Table 2 shows the fractional changes of resistance ($\Delta\rho/\rho_0$) observed at the temperatures of liquid nitrogen and liquid helium, and Figure 3 gives these results in the form of a "Köhler diagram" or log-log plot of $\Delta\rho/\rho_0$ against the ratio H/r_T of the magnetic field to the reduced resistance. Included in Figure 3 are some results of Schulze[7] for the magneto-resistance of two specimens of rhodium, one of higher and one of lower purity than that used in the present investigation.

Table 2. Fractional change ($\Delta\rho/\rho_0$) in electrical resistance of rhodium in a transverse magnetic field at 77°K and 4·2°K

77°K		4·2°K	
H (kilogauss)	$\Delta\rho/\rho_0 \times 10^3$	H (kilogauss)	$\Delta\rho/\rho_0$
7·35	1·52	3·4	0·0326
9·46	2·44	4·8	0·0586
11·50	3·83	6·2	0·0841
13·45	5·66	7·6	0·1101
18·35	9·58	8·8	0·1358
20·75	13·04	12·1	0·2088
22·70	14·93	15·2	0·2724
		18·6	0·3527
		20·6	0·3996

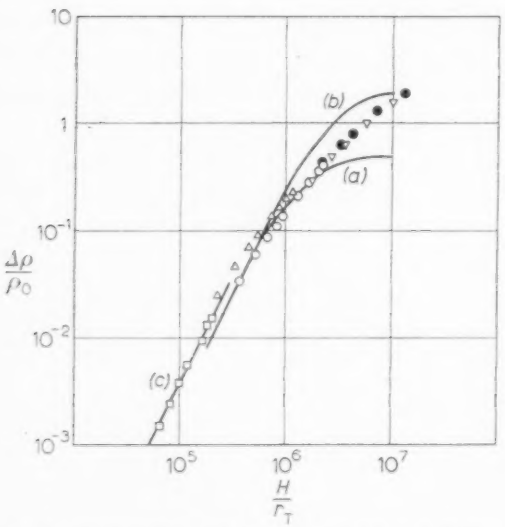


FIG. 3. Fractional change in resistance ($\Delta\rho/\rho_0$) of rhodium as a function of the ratio (H/r_T) of the magnetic field in gauss to the reduced resistance. The solid curves are those given by the two-band model using the values of the parameters given in the text.

○ Present work 4·2°K, $r_T = 9{\cdot}32\times10^{-3}$
⊡ Present work 77·0°K, $r_T = 113{\cdot}70\times10^{-3}$
● Schulze (Rh II) 4·2°K, $r_T = 2{\cdot}90\times10^{-3}$
▽ Schulze (Rh II) 20·4°K, $r_T = 3{\cdot}58\times10^{-3}$
△ Schulze (Rh I) 20·4°K, $r_T = 35{\cdot}30\times10^{-3}$

4. DISCUSSION

4.1. *The Hall Effect*

The field dependence of the Hall effect has been discussed in terms of two band models by JONES[8] and SONDHEIMER,[9] and general expressions have been derived. The Hall coefficient is conveniently given by the expression

$$R = \frac{-B}{Nec}\left[\frac{1 + A(1/Nec)^2(H/\rho_0)^2/B}{1+(n_1-n_2)A(1/Nec)^2(H/\rho_0)^2}\right]$$

where

$$B = p^2/n_1 - (1-p)^2/n_2$$

and

$$A = p^2(1-p)^2(n_1-n_2)/(n_1n_2)^2,$$

n_1 and n_2 being the respective numbers *per atom* of electrons in one band and holes in the other, p the fraction of the total current carried by the n_1 electrons, N the number of atoms per unit volume, ρ_0 the electrical resistivity in zero magnetic field and H the magnetic field; e and c have their normal significance. This expression can be made to fit the experimental results quite closely. In Figure 2, the solid curves are drawn using $+3{\cdot}34$ for A and $+4{\cdot}08$ for (n_1-n_2); ρ_0 was obtained from the values of r_T given in Table 1 using the measured value of $4{\cdot}3\times10^{-6}$ ohm cm for the resistivity at 273°K. The quantity B (which governs the low field Hall coefficient) varies slightly with temperature and the values used in constructing the theoretical curves are those obtained from the R_0 values shown in Table 1. The origin of this variation must presumably be a

slight temperature dependence of one or more of the parameters n_1, n_2 and p, but all three curves of Figure 2 have been constructed with the same values of (n_1-n_2) and A.

Using the above values for A and (n_1-n_2) and the 20·4°K value for B, one obtains for the parameters involved

$$n_1 = 4{\cdot}25; \quad n_2 = 0{\cdot}17; \quad \text{and} \quad p = 0{\cdot}675.$$

It would probably be unwise to attempt to attach any detailed significance to the exact values thus obtained, but if the two-band model has any validity for this metal the following remarks seem justified: first, that the high mobility carriers (the sign of which controls that of the low-field Hall coefficient) must be holes; second, that comparatively small values of either n_1 or n_2 or both must be responsible for the marked field dependence; and third, that for the Hall coefficient to show the observed reversal of sign the number of holes must, since B is known to be negative, be smaller than that of electrons.

It is difficult to relate these observations at all directly to the electronic structure of rhodium. As a transition metal a two-band model would be expected to be appropriate for it, empty states existing in a 4*d*-band and occupied ones in a band built up mainly from 5*s* and 5*p* states; but with a total of 9 outer electrons the number of *d*-band holes must always exceed by one that of the (*sp*) electrons. Furthermore *d*-band holes might be expected, in general, to have a lower mobility than the (*sp*) electrons, whereas the above values of n_1, n_2 and p imply a very much higher relative mobility for the n_2 holes. It is possible, however, that the number of empty *d*-states in metallic rhodium is sufficiently large for the (*sp*) electrons to lie in two Brillouin zones and for the *d*-band electrons to show conduction of negative sign, the positive Hall coefficient arising (as it does in cadmium) from the positive holes in the first Brillouin zone containing (*sp*) electrons. The actual numbers given above for n_1 and n_2 would then have no significance for such a three-band situation. This situation could arise if the *d*-band in rhodium lay at appreciably higher energy (relative to the (*sp*) band) than it seems to in palladium. Such a relative movement of the *d* and (*sp*) bands as the atomic number is reduced has been discussed[2] for the end members of the first transition group; but there the initial movement seems much less, that between nickel and cobalt bringing about an increase from 0·6 to about 0·75 in the number of (*sp*) electrons for the face-centred cubic phases of nickel and cobalt. (Recent work on iron-cobalt alloys[10] suggests a rather more rapid increase in this number on passing back to iron.)

The resultant difference between rhodium and cobalt is not, however, surprising, for rhodium and ruthenium are known to exhibit higher valencies in chemical combination than cobalt and iron. On the other hand, while spectroscopic data for the free atoms and ions[11] show the general relative movement of *d* and *s* levels expected, the character of this movement does not seem to differ greatly for the first and second transition groups.

4.2. *Magneto-Resistance*

The general expression for the change of resistance $(\Delta\rho)$ in a transverse magnetic field from the zero field resistance (ρ_0) is given[12] by

$$\Delta\rho/\rho_0 = \frac{C(1/Nec)^2(H/\rho_0)^2}{1+(n_1-n_2)A(1/Nec)^2(H/\rho_0)^2}$$

where A has the same significance as in (1) and

$$C = [p/n_1+(1-p)/n_2]^2 p(1-p).$$

The solid curves in Figure 3 are given by expressions of this type. For (a) $C = 6{\cdot}8$ and $A = 3{\cdot}34$, the value used in fitting the Hall effect data, for (b) $C = 6{\cdot}8$ and $A = 0{\cdot}82$, and for (c) $C = 9{\cdot}75$ and $A = 0$. These values for C differ greatly from that of 0·89 given by taking the former values of n_1, n_2 and p. It is evident from Figure 3 that a factor of the form $(1+K(H/\rho)^2)^{-1}$ gives too rapid a change of slope with H/r_T, whatever the value of K. It is also obvious that, contrary to "Köhler's rule", the temperature dependence is not governed only by that of ρ_0; the change in the low field magneto-resistance coefficient between 77°K and 4·2°K is very much greater than the corresponding change in the low field Hall coefficient. This observation is in agreement with the stronger dependence on p shown by C as compared with

B, if the temperature variation of B is correctly to be ascribed to variations in p.

Acknowledgements—The authors' thanks are due to Dr. F. E. Allison, Dr. S. A. Friedberg, and Dr. E. M. Pugh for helpful co-operation and discussion, and to Dr. E. C. Rhodes, Mr. J. R. Knight and the Mond Nickel Company (Precious Metals Division) for the provision of the rhodium sample. Grateful acknowledgement is made to the U.S. International Co-operation Administration for a fellowship held by one of the authors (B.R.C.), to the U.S. Office of Naval Research for financial support, and to the Ford Motor Company, Scientific Laboratory for hospitality during the later stages of this work.

REFERENCES

1. Taylor J. C. and Coles B. R. *Phys. Rev.* **102,** 27 (1956).
2. Hume-Rothery W. and Coles, B. R. *Adv. in Phys.* **3,** 149 (1954).
3. Coles B. R. *Phys. Rev.* **101,** 1254 (1956).
4. Schindler A. I. and Pugh E. M. *Phys. Rev.* **89,** 295 (1953).
5. Foner S. and Pugh E. M. *Phys. Rev.* **91,** 20 (1953).
6. Gehloff P. O. and Justi E. *Z. Naturforsch.* **A4,** 561 (1949).
7. Schulze R. *Phys. Z.* **42,** 297 (1941).
8. Jones H. *Proc. Roy. Soc.* **A155,** 653 (1936).
9. Sondheimer E. H. *Proc. Roy. Soc.* **A193,** 484 (1948).
10. Coles B. R. and Bitler, W. R. *Phil. Mag.* (8) **1,** 477 (1956).
11. Catalan M. A., Rohrlich F., and Shenstone A. G. *Proc. Roy. Soc.* **A221,** 421 (1954).
12. Sondheimer E. H. and Wilson A. H. *Proc. Roy. Soc.* **A190,** 435 (1947).

VOL. 1 1956/5

J. Phys. Chem. Solids. Pergamon Press 1957. Vol. 1. pp. 275–278.

THE EFFECT OF GAMMA IRRADIATION ON THE YOUNG'S MODULUS OF COPPER

D. O. THOMPSON AND D. K. HOLMES

Solid State Division, Oak Ridge National Laboratory

(*Received* 12 *October* 1956)

Abstract—Copper single crystals have been subjected to irradiation with 1·2 MeV Co^{60} gamma rays at room temperature. This irradiation causes an increase in the Young's modulus and a decrease in the internal friction. These effects are quantitatively consistent with an interpretation on the basis of the pinning of dislocation lines by lattice defects which migrate to the dislocation line from the point at which they are originally produced through the Compton and photoelectrons resulting from the incident gamma rays. Comparison with previous work shows that for each primary displaced lattice atom only one eightieth as many pinning points result from the gamma ray bombardment as from bombardment with fast neutrons; this result is in satisfactory agreement with calculations of higher order displacements in the two cases.

RECENT work[1, 2] has shown that irradiation of copper single crystals with fast neutrons results in an increase in the Young's Modulus in a manner which is consistent with an interpretation in terms of the pinning of dislocation line segments by the resultant defects from the neutron bombardment (presumably vacancies, interstitials, or clusters of these). The analysis of the results of neutron damage does not lead to unique values for dislocation line densities and initial length distribution of line segments because of the lack of information as to the number of pinning points produced per elastic collision by a fast neutron. In order to obtain a value for this number it would be necessary to know the precise mechanism for the pinning and also to know the number and state of clustering of defects produced per neutron collision. Some of the uncertainty in the number of defects produced (and possibly in the state of clustering) might be removed by the use, as the bombarding particle, of electrons with energies just above the threshold for the displacement of lattice atoms. However, the short range of such electrons introduces experimental difficulties in the measurement of modulus.

A more fruitful approach is found in the work of DUGDALE[3] who showed that irradiation by gamma rays can produce measurable effects in metals. The lattice atoms are actually displaced by electrons produced in the lattice by the gamma rays through the Compton effect and the photoelectric effect. A further demonstration of this effect has recently been obtained at this laboratory by the bombardment of germanium with Co^{60} gamma rays.[4] The advantages in the use of gamma rays are:

1. The production of electrons, and thus the damage, is practically uniform over the volume of specimens whose dimensions are small in comparison with a relaxation length for gammas of the particular energy chosen.
2. For gammas of sufficiently low energy, the primary displaced atoms do not have enough energy to produce secondary displacements. Consequently, no more than one vacancy-interstitial pair may be counted per elastic collision by an electron, and each pair produced will, on the average, be many thousands of lattice spacings removed from the nearest neighboring pair.

As a disadvantage it must be mentioned that the displacement cross-sections are low, for low-energy gammas ($\sim$ 1 Mev), so that, with available sources of gamma rays, the results of the irradiation can be observed only on very sensitive properties

of the material. From the fact that neutron fluxes of the order of only 10^{12} neutrons/cm²/sec are needed to produce easily detectable changes,[2] it was judged that the dislocation contribution to the Young's modulus is sufficiently sensitive for the observation of gamma ray effects.

The measurements were performed as previously described for neutron bombardment[2] near a 900 Curie Co^{60} source, in such a position as to provide a flux of about 4×10^{10} gamma/cm²/sec (1·17 Mev and 1·33 Mev in equal numbers). Crystal 5A, which had previously been irradiated to saturation in a neutron flux[2] and subsequently annealed, was used for the irradiation. The bombardment and measurements were carried out at about 25°C. Data for the Young's modulus are shown in Fig. 1. It was found that the results follow very well the prediction of the random

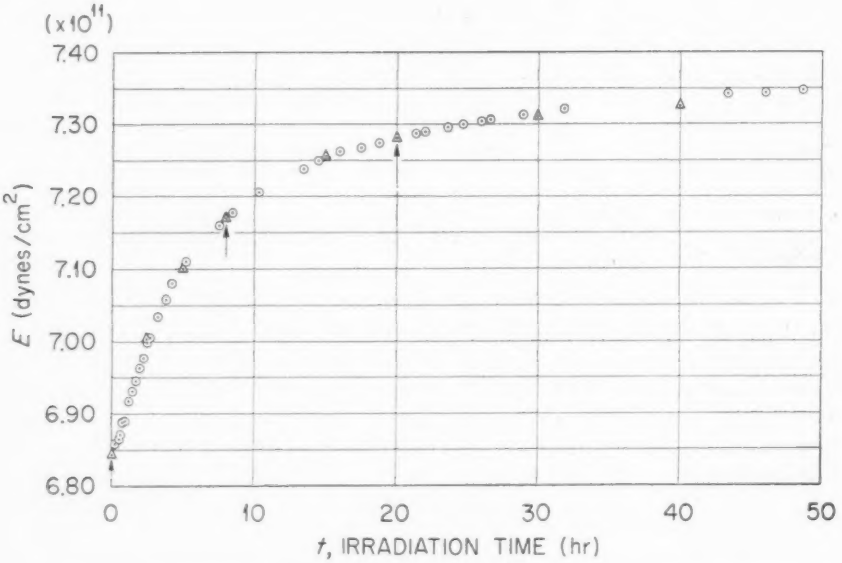


FIG. 1. The variation of the Young's modulus of a copper single crystal with time of exposure to gamma rays.

⊙ Data points.
△ Points calculated from eq. 1 of the text.
↑ Fitting points used to obtain eq. 1.

dislocation pinning theory[2] as shown by the points on Fig. 1 which are enclosed in triangles; these are calculated from the equation (for $E(t)$ = the Young's modulus as a function of irradiation time)

$$\textit{Gammas}: \quad \frac{1}{E(t)} = \frac{1}{7{\cdot}348\times10^{11}} + \frac{1{\cdot}0046\times10^{-13}}{(1+0{\cdot}0905t)^2} \qquad (1)$$

The final value of the modulus agreed very closely with that resulting from the neutron bombardment, as would be expected on the dislocation pinning model, since the saturation value should be just the modulus of the nearly perfect copper lattice. The results of the neutron bombardment on the same crystal are given by

$$\textit{Neutrons}: \quad \frac{1}{E(t)} = \frac{1}{7{\cdot}348\times10^{11}} + \frac{2{\cdot}896\times10^{-13}}{(1+0{\cdot}0108t)^2} \qquad (2)$$

These equations are of the form previously obtained,[2] i.e.

$$\frac{1}{E(t)} = \frac{1}{E_e} + \frac{\beta^2 L_0 l_0^2}{\mu(1+\theta l_0 t)^2} \tag{3}$$

in which

E_e = the purely elastic modulus of the crystal.
β^2 = orientation factor.
L_0 = the total (original) length of dislocation line per unit volume.
l_0 = the average (original) length of dislocation line segments.
μ = the shear modulus.

In general, as indicated in reference (2), the quantity "θ" has the form

$$\theta = \frac{nN\sigma_d\phi}{L_0} \tag{4}$$

where

φ = flux of bombarding particles.
N = the number of lattice atoms per cm^3.
σ_d = the atomic cross-section for displacement of a lattice atom by the incident particle.
n = the number of pinning points on dislocation line segments produced per initial displaced atom.

In order to compare the pinning effectiveness for the two types of bombarding particles, it is necessary to compare the experimental values obtained for θl_0 in the two cases. This ratio is

$$\frac{(\theta l_0)_\gamma}{(\theta l_0)_n} = \frac{0{\cdot}0905}{0{\cdot}0108} \cong 8{\cdot}38 = \frac{n_\gamma \phi_\gamma \sigma_\gamma (l_0)_\gamma}{n_n \phi_n \sigma_n (l_0)_n} \tag{5}$$

in which n_γ/n_n is the desired quantity. Since the initial modulus in the case of the gamma irradiation was slightly lower than in the neutron runs, it appears that the ratio $(l_0)_\gamma/(l_0)_n$ is not equal to one. This must be attributed to the slight handling of the crystal necessary to introduce it into the gamma field. If the reasonable assumption is made that the dislocation line density, L_0, of the crystal is unaltered by the slight handling, then the reduction in the modulus must be attributed to a change in the average loop length, l_0. The required lengthening in l_0 on handling might, for example, be attributed to the tearing away of the dislocation lines from a few impurity atom pinning points. From other experience of this type it is expected that, at room temperature, the restoration of these pinning points would be very slow in comparison with the rate of pinning point introduction by the gamma bombardment. If this interpretation is chosen, a comparison of the values in the numerators of the second terms on the right hand sides of equations (1) and (2) with the theoretical result of equation (3), gives a value for the ratio of the average original loop length in the gamma case to that in the neutron case of 1·7. Using the measured values for the neutron and gamma fluxes, $3{\cdot}1\times10^6$ and 4×10^{10}, respectively (these values are not expected to be accurate to better than 20 per cent) and the loop length obtained above, the ratio of the average number of pinning points per initially displaced lattice atom for the two radiations is

$$\frac{n_n}{n_\gamma} \cong 2{,}620\frac{\sigma_\gamma}{\sigma_n} \tag{6}$$

The cross-sections of equation 6 are the atomic cross-sections for displacement by the particular radiation; in the case of the neutrons, this is just the neutron scattering cross-section, which, for copper, is about four barns. The cross-section per atom for displacement by a gamma ray may be calculated using, for example, the information given by BETHE and ASHKIN,[5] as 0·120 barns, if it is assumed that the displacement threshold energy for a lattice atom in copper is 25 ev. This cross-section is primarily (95 per cent) due to the energetic Comption electrons produced by the gamma rays. There is a small contribution due to the photoelectrons produced by the gammas; the photoelectric recoils of sufficient energy to displace an atom occur in negligible quantities for gammas of these energies. (The analogous results for germanium are incorrectly given in reference 4. The displacement cross-section is 0·123 barns, essentially the same as for copper and is also primarily due to the Compton process.) The final result, then, for the ratio of equation 6 is

$$\frac{n_n}{n_\gamma} \cong 80 \tag{7}$$

In consideration of this result it must be pointed out immediately that due to difficulties to be discussed below, the numerical value given in equation 7 is not expected to be accurate to better than

a factor of two or three. From considerations of the type presented by SNYDER and NEUFELD[6] it may be estimated that about 110 displacements (vacancy-interstitial pairs) should be produced (neglecting annealing) per fast neutron in the reactor spectrum.* If it is supposed that only one displacement (vacancy–interstitial pair) is produced per initial displaced atom in the gamma ray case, then the figure of equation 7 is of about the expected magnitude.

However, the matter is not really so simple; other experiments have indicated that the theoretical figure quoted above for displaced atoms per reactor fast neutron is too high by a factor of five. Further, the results from the bombardment of germanium with gamma rays[4] indicate an atomic displacement cross-section of 0·015 barns per incident gamma as compared with a value of 0·123 barns calculated on the same basis as used for the copper. Thus, it seems that due either to annealing at the temperature of measurement (25°C) or to incorrect assumptions in the calculation, the effective displacement cross-section is only about 10 per cent of the calculated value. If the same percentage effectiveness is allowed in copper, and if each vacancy and each interstitial are effective in pinning dislocation lines, then n might be estimated as about 0·2; on this basis n_n, from the result of equation 7, would be about 16.0. On this basis, the average original dislocation loop length (in the neutron case) would be about 3×10^{-5} cm and the dislocation line density would be about 10^8 cm^{-2}.

These figures seem entirely reasonable; however, it must be remembered that there are great uncertainties in this analysis. One of the most important of these is the nature of the threshold for atomic displacements. Recent experiments and interpretations[7] have indicated that the threshold may not be sharp, and that displacements are possible with transferred energies much less than the 25 ev as assumed in obtaining the gamma ray displacement cross-section.

* Actually, the figure 110 is based on the spectrum at the center of the X–10 Graphite Reactor, whereas the neutron experiments were performed outside a two-foot graphite reflector; it would be expected that the slowing-down of the neutrons in the graphite would reduce their effectiveness in producing lattice displacements, but no estimate of this reduction is available.

REFERENCES

1. THOMPSON D. O. and HOLMES D. K. *J. Appl. Phys.* **27,** 191 (1956).
2. THOMPSON D. O. and HOLMES D. K. *J. Appl. Phys.* **27,** 713 (1956).
3. DUGDALE, R. A. *Report of Bristol Conference on Defects in Crystalline Solids.* London, Physical Society (1955).
4. CLELAND J. W., CRAWFORD J. H., and HOLMES D. K. *Phys. Rev.* **102,** 722 (1956).
5. BETHE H. A. and ASHKIN J. *Experimental Nuclear Physics.* ed. E. Segre, (John Wiley & Sons) New York, (1953) Vol. 1, p. 166.
6. SNYDER W. and NEUFELD J. *Phys. Rev.* **97,** 1636 (1955).
7. SNYDER W. and NEUFELD J. *Phys. Rev.* **103,** 862 (1956).

VOL. 1 1956/57

J. Phys. Chem. Solids. Pergamon Press 1957. Vol. 1. pp. 279–283.

ZUR ELEKTRONENTHEORIE METALLISCHER UND HALBLEITENDER STRUKTUREN

U. DEHLINGER

Institut für theoretische und angewandte Physik der Technischen Hochschule Stuttgart und Max-Planck-Institut für Metallforschung, Stuttgart

(*Received* 2 *October* 1956)

Zusammenfassung — Bei den Gittern mit metallischer Leitfähigkeit — Elementen und Verbindungen — ist zu unterscheiden zwischen gerichteter (kovalenter) Bindung durch Kombinationen von Atomfunktionen und zwischen ungerichteter Bindung, für die die Gesetze des nahezu freien Elektronengases angewandt werden können. Die erstere wird mit wachsender Zahl der Bindungselektronen und abnehmender Hauptquantenzahl gegenüber der ungerichteten bevorzugt. In den Gittern mit gerichteter Bindung ist also die Reihenfolge der Bänder durch die Kombination von Atomfunktionen und damit durch Atomterme bestimmt. Die Bänder spalten in bindende und lockernde Teilbänder auf. So ist das Leitfähigkeitsband der diamantartigen Halbleiter als lockerndes Teilband der *sp*-Funktion aufzufassen.

Abstract—In the elements and compounds with metallic conductivity directed (covalent) binding by combinations of atomic functions is very different from undirected binding, for which the laws of nearly free electrons may be applied. The first type is preferred in lattices with more than two binding electrons per atom, if the principal quantum number is not loo large. In the periodic system one sees a sharp boundary between the two types. In Ga and in the transition metals both types are formed side by side. In directed binding the sequence of bands is to be fixed by the atomic functions and therefore mostly by the atomic terms. The bands split up in binding and unbinding half-bands. The conductivity band of diamant-like semiconductors has to be interpreted as the unbinding half-band of the (*sp*)-function.

DER Verfasser hat mit einer Reihe von Mitarbeitern versucht, die Strukturen von Elementen und intermetallischen Verbindungen elektronentheoretisch zu deuten. Ausgehend von Diamant wurde das metallisch leitende β-Sn (weisses Zinn) behandelt.[1] Zunächst zeigte die gruppentheoretische Analyse, dass dieselben "gerichteten" *sp*-Atomfunktionen wie im Diamantgitter in Betracht kommen. Aus ihnen werden nach Bloch die noch von k abhängenden Einelektronen-Gitterfunktionen (molecular orbitals) gebildet, auf die eine Störungsrechnung angewandt wird. Die Störungsenergie berechnet sich aus den für die nächsten, übernächsten usw. Nachbarn gebildeten Resonanz- oder Überlappungs-Integralen, wobei die Winkel zwischen den Vorzugsrichtungen der Atomfunktionen und den Valenzstrichen wesentlichen Einfluss haben. So erhält man eine Abschätzung für die Breite und die Lage der Bänder in Bezug auf das Niveau des freien Atoms. Wie beim Diamant spaltet jedes Band in ein bindendes und ein lockerndes Teilband auf. Es ist dies immer dann zu erwarten, wenn ein Symmetriezentrum in der Mitte zwischen zwei Atomlagen, nicht aber in diesen selbst liegt, wenn also die kleinste Grundzelle mindestens zwei Atome enthält.[2] Beim Diamant sind die beiden Teilbänder (meist Valenz- und Leitungsband genannt) um 7 eV getrennt, mit steigender Hauptquantenzahl nimmt der Abstand gesetzmässig ab, weil des für ihn entscheidende Resonanzintegral entsprechend abnimmt.[1]* Für α-Sn ist der Abstand zu etwa 0,1 eV gemessen, im β-Sn

* Die azimutale Verteilung der einzelnen *s*–oder *p*-Funktionen hängt bekanntlich nicht von der Hauptquantenzahl ab. Anders ist es bei den Linearkombinationen von *s*– und *p*– bezw. *d*-Funktionen. Da diese verschiedene Radialanteile besitzen, werden die Kombinationen mit steigender Hauptquantenzahl *s*-ähnlicher.

sind die Bänder wegen der erhöhten Koordinationszahl breiter und überlappen daher, was zu der metallischen Leitfähigkeit dieser Modifikation führt. In derselben Weise konnten As, Sb, Bi behandelt werden,[3] wobei auch hier die Hauptquantenzahl starken Einfluss auf die Breite und die Lage der Bänder besitzt, und es ist kein Zweifel, dass alle Strukturen von Elementen hierher gehören, welche die obenerwähnten Symmetrieverhältnisse besitzen. In grundsätzlich gleicher Weise wurden schon vorher die *d*-Schalen der Übergangsmetalle behandelt,[2, 4] die ja wegen ihrer Unabgeschlossenheit nicht die volle Symmetrie des Gitters besitzen können.

Die Vorzugsrichtungen der *d*-Funktionen, die je nach der Auffüllung der Bänder die Allotropie der Übergangsmetalle bestimmen, sind auch in den Verbindungen von Übergangsmetallen für die Struktur entscheidend, z.B. im NiAs-Typ,[5] im Fe_3C[5] und $MoSi_2$.[6]* In allen diesen Gittern fallen die Winkel zwischen den Bindungsstrichen zu den nächsten, zum Teil auch übernächsten Nachbarn genau oder nahezu zusammen mit den Winkeln zwischen den Vorzugsrichtungen der bindenden *d*- und *sp*-Funktionen.

Die Verbindungen, die eine solche "Kovalent-metallische" Bindung besitzen, treten selbstverständlich nur bei ganz bestimmten Elektronenkombinationen auf. Zum Beispiel erscheint der Typ des Fe_3C nur noch bei Ni_3C, der des $MoSi_2$ nur noch bei WSi_2, der NiAs-Typ tritt nur dann auf, wenn die von Zintl genau studierten elektronischen Bedingungen vorhanden sind.[2]

Da die *sp*-Funktionen im Gegensatz zu den *d*- und *sd*-Funktionen kein Symmetriezentrum besitzen können, gilt der Satz: Wird die gerichtete Bindung eines Atoms allein durch kombinierte *s*- und *p*-Funktionen erzeugt, dann tragen diese Atome kein Symmetriezentrum.[5] Wie die Erfahrung zeigt, ist dieser Satz in allen obengenannten Fällen erfüllt.

Das bisher Gesagte, in dem vor allem der Einfluss der Hauptquantenzahl deutlich hervortritt, kann als empirischer Beweis dafür angesehen werden, dass bei der Darstellung der Gitter-Eigenfunktionen (molecular orbitals) nach BLOCH

$$\Psi = e^{2\pi i k r} u_k(r)$$

$$u_k(r) = \sum_i e^{-2\pi i k(r-r_i)} v(r-r_i) \qquad (1)$$

die den wirklichen Hauptquantenzahlen zugeordneten *s*-, *p*- und *d*-Funktionen bezw. ihre Linearkombinationen entscheidend sind, d.h. in einer Entwicklung der v bei der Berechnung der Term-Energie gegenüber höheren Gliedern den überwiegenden Beitrag geben. Wenn man also die v allein durch die genannten Kombinationen von Atomfunktionen ausdrückt, ist das für die Gitter mit gerichteter Bindung eine bessere Annäherung der Eigenfunktionen als in Anbetracht des kleinen Atom-Abstands zu erwarten wäre. Die energetische Reihenfolge der Terme des Gitters ist wesentlich bestimmt durch die Reihenfolge der Atomterme, die Bänder können also durch die Hauptquantenzahlen und die Angabe der *s*-, *p*- und *d*-Kombinationen gekennzeichnet werden. Des "Valenzband" ist identisch mit dem bindenden, das "Leitungsband" mit dem lockernden Teilband dieser Bänder.

Der Grund für diese empirische Tatsache ist offensichtlich darin zu suchen, dass die gerichteten Atomfunktionen durch ein genügend stark von der Kugelsymmetrie abweichendes, störendes Feld, in das ihre azimutale Verteilung passt, energetisch bevorzugt werden. Andererseits wird dieses störende Feld durch die mit den bindenden Elektronen besetzten gerichteten Funktionen selbst aufrecht erhalten. Es handelt sich also zweifellos um ein Problem des selfconsistent field, wie es für Gitter bisher noch nicht behandelt wurde. In Übereinstimmung damit steht die empirische Tatsache, dass Gitter mit gerichteter Bindung nur von mindestens dreiwertigen Elementen gebildet weden, und dass ihre Stabilität gegenüber den Gittern mit nichtgerichteter Bindung mit steigender Wertigkeit, aber abnehmender Hauptquantenzahl wächst. Mit abnehmender Hauptquantenzahl wird nämlich die Schärfe der Vorzugsrichtungen in den gerichteten Atomfunktionen grösser, und damit die Abweichung von der Kugelsymmetrie stärker.[1, 2]

Bei Vergrösserung des Atomabstands wird die gerichtete Bindung gegenüber der ungerichteten bevorzugt, da sie ja bei nahezu freien Atomen sicher vorhanden ist. Als Beispiel dafür mögen Bindungsverhältnisse in PbS dienen, das im NaCl-Typ kristallisiert[7] und ein Halbleiter ist.

* In dieser Arbeit wurde auch die gerichtete Bindung im Ga-Gitter berechnet.

VOL. 1 1956/5

Das NaCl-Gitter verlangt, wie man aus der Legierungschemie weiss,[2] weitgehende Heteropolarität, man kann also davon ausgehen, dass Pb und S zweifach geladen sind. Das Ion Pb^{++} hat noch zwei Aussenelektronen. Sie besetzen offensichtlich eine s-Funktion, die ein schmales und nicht überlappendes Band bildet. Die Halbleitereigenschaft kann nur dadurch zu Stande kommen, dass zwischen diesem *s*-Band und dem darauffolgenden unbesetzten Band eine kleine Lücke ist. Das letztere Band wird gebildet von drei nach den Würfelkanten gerichteten, miteinander entarteten *p*-Atomfunktionen,* vielleicht auch von nach den Würfeldiagonalen gerichteten *d*-Funktionen. Wenn der Abstand der Pb-Atome so klein wäre, wie im metallischen Pb, dann wären die *p*- und *d*-Funktionen mit den *s*-Funktionen zusammen entartet und man hätte ungerichtete Bindung, wie sie in der Tat für die Bindung zwischen den nächsten Nachbarn des NaCl-Typs anzunehmen ist.

Alles bisher Gesagte gilt zweifellos nicht für die Elemente mit innenzentriertem, flächenzentriertem und hexagonal dichtgepacktem Gitter, soweit sie abgeschlossene *d*-Schalen besitzen. Hier sind keine eindeutigen Beziehungen zwischen Vorzugsrichtungen der Atomfunktionen und Bindungsstrichen zu finden, es liegt also eine nicht gerichtete Bindung vor. Für die einwertigen Metalle hat u.a. eine kritische Sichtung der optischen und thermischen Messungen durch A. Seeger und H. Kronmüller[8] gezeigt, dass man das Verhalten der Bindungselektronen mit guter Übereinstimmung durch eine einzige effektive Masse wiedergeben kann, die nicht richtungsabhängig ist. Unter den zwei- und dreiwertigen Metallen befolgen die Strukturen des hexagonalen Cd und des tetragonalen Jn die Theorie von Jones, insofern als sie bei Veränderung ihrer Elektronenzahl durch kleine Zusätze ihr Achsenverhältnis so ändern, wie es durch die Überragungen der nahezu kugelförmigen Fermikörper über die erste Brillouinzone gefordert wird.[2]

Diese empirischen Fakten zeigen, dass bei den Metallen mit nicht gerichteter Bindung die Entwicklung der Gittelfunktionen in eine Fourierreihe

$$U_k(r) = \sum_g a_{kg} e^{2\pi i g r}$$

insofern vorzuziehen ist, als in ihr die ersten Glieder überwiegen, während bei einer Entwicklung von der Form (1) sehr viele Glieder mit nahezu gleichem Einfluss notwendig sind. Die Terme des Leitfähigkeitsbands ordnen sich also nach zunehmenden Werten von $(k+g)$ und die Termabstände der Funktionen des freien Atoms, aus denen sie hervorgehen, sind infolge des von den umgebenden Atomen herrührenden Störungsfeldes unmerklich klein geworden. Besonders zu behandeln sind in diesem Zusammenhang die Elemente Mg und Al. Bei Mg tritt die bei den ebenfalls zweiwertigen Zn und Cd gefundene,

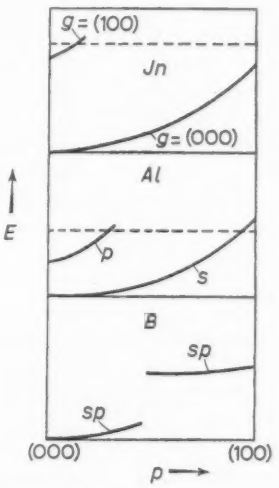


Fig. 1. Schematischer Verlauf der Energie E mit K bei dreiwertigen Elementen verschiedener Hauptquantenzahl. Die erste Brillouinzone wird bei In überragt, dagegen bei Al nicht.

für einen nahezu kugelförmigen Fermikörper charakteristische Gitterkonstantenabhängigkeit von der Valenzelektronenzahl nicht auf.[2] Dasselbe zeigt ein Vergleich der dreiwertigen Elemente Al und Zn; insbesondere hat Al niemals ein tetra-

* Solche *p*-Funktionen treten nach U. Dehlinger, H. Schenk und W. Weser, Z. *Metallkunde* **46**, (1955) 647 auch im As-Gitter auf, solange dieses als einfach kubisch angesehen werden können. Bei der kleinen rhomboedrischen Verzerrung, die tatsächlich vorhanden ist, kombinieren *s*- und *p*-Funktionen ein wenig.

gonales Gitter. Somit ist empirisch sicher, dass in Mg und Al der Fermikörper wesentlich stärker von der Kugelform abweicht als in den gleichwertigen Elementen höherer Hauptquantenzahl. Die energetische Reihenfolge der Terme wird also etwa dem in Fig. 1 gezeichnenten Schema folgen, daher werden die für Zn bezw. In zu erwartenden Überragungen des Fermikörpers hier noch nicht eintreten. Es zeigt sich also empirisch, dass die bei der vollständig ungerichteten Bindung auftretende vollständige Entartung der Atomterme bei kleiner Hauptquantenzahl, wo ja der Abstand der Atomterme grösser wird, noch nicht vorhanden ist.

Auch unter den (metallischen und nicht metallischen) Verbindungen tritt die ungerichtete Bindung in wichtigen Fällen auf. Zur Unterscheidung von der gerichteten Bindung kann hier vielfach ein neues Kriterium dienen, nämlich der theoretisch einleuchtende Umstand, dass die ungerichtete Bindung (vor allem, wenn sie metallisch ist, also keine aufgefüllten Bänder verlangt) nicht so wie die gerichtete nur auf ganz bestimmte Kombinationene von Elektronen beschränkt ist. So besitzen nicht gerichtete Bindung die Gittertypen des $CuAl_2$, der Flusspatstruktur des $AuAl_2$, sowie die sog. Lavesphasen vom Typ $MgCu_2$, $MgZn_2$ und $MgNi_2$, weiterhin die Cäsiumchloridstruktur des CuZn oder LiAg und die Gitter mit der Struktur des AuCu und $AuCu_3$. In allen diesen Fällen wird der Typ gebildet von Elementkombinationen,[9] die über fast das ganze periodische System gestreut sind, und deren Gesetzmässigkeiten allen aus der sonstigen Chemie bekannten Gesichtspunkten zu widersprechen scheinen.

In den genannten Fällen genügen die Gesetze des nahezu freien Elektronengases, um das Zustandekommen der Gittertypen wenigstens qualitativ zu erklären. Und zwar handelt es sich um folgende Forderungen an die Symmetrie und Koordinationszahl der Gitter: 1) Die mittlere Koordinationszahl muss so gross sein, als sich mit den Atomradienverhältnissen vereinbaren lässt. Diese Forderung allein bestimmt u.a. die Gitter der Lavesphasen.[10] 2) Werden Elemente verschiedener Wertigkeit miteinander legiert, so tritt nach J. FRIEDEL[11] eine teilweise Heteropolarität auf, die in der Nähe von 50 Atom% ihr Maximum hat. Ebenso kann eine Heteropolarität dadurch entstehen, dass die stark unedlen Alkali- und Erdalkalimetalle sich teilweise positiv laden. Diese Heteropolarität begünstigt im flächenzentrierten Gitter eine regelmässige Atomverteilung vom Typ AuCu und $AuCu_3$, führt aber, wenn sie stärker ist, in Konkurrenz zur Forderung 1) zu einem innenzentrierten CsCl-Typ[2] sowie zum Flusspat und $CuAl_2$-Typ,[13] der ebenfalls die Koordinationszahl acht zwischen A- und B-Atomen hat. 3) durch die bei nahezu kugelförmigem Fermikörper entstehenden Überragungen von Brillouinzonen werden die Existenzgebiete von flächenzentrierten und innenzentrierten Gittern beeinflusst, ausserdem entstehen aus dem CsCl-Typ die γ-Gitter. Dagegen hat bei gerichteter Bindung die Auffüllung von Brillouinzonen keinen Einfluss auf die Struktur. Ähnlich wie bei der gerichteten Bindung scheint auch bei der ungerichteten kein grundsätzlicher Unterschied zwischen Fällen mit und ohne metallische Leitfähigkeit zu bestehen. Insbesondere gehen die oben genannten Phasen mit teilweiser Heteropolarität ohne deutliche Grenze in der Struktur in solche mit voller salzartiger Bindung und valenzmässiger Zusammensetzung über.* Insgesamt dürfte die Unterscheidung zwischen gerichteter und nicht gerichteter Bindung es ohne weitere Schwierigkeiten ermöglichen, die jedem der beiden Fälle eigentümlichen Gesetze darzustellen und damit das Gesamtgebiet der metallischen Phasen sowie vieler Halbleiter elektronentheoretisch zu ordnen.

VOL. 1 1956/57

LITERATUR

1. DELINGER U. und SCHENK H. *Z. Physik* **126,** 344 (1953).
2. DEHLINGER U. *Theoretische Metallkunde* Berlin-Göttingen-Heidelberg (1955).
3. DEHLINGER U., SCHENK H., und WESER W. *Z. Metallkde.* **46,** 647 (1955).
4. BADER F., GANZHORN K., und DEHLINGER U. *Z. Physik* **137,** 190 (1954).
5. SCHENK H. und DEHLINGER U. *Z. Naturf.* **11a,** 327 (1956).
6. SCHENK H. und DEHLINGER U. *Acta Met.* **4,** 8 (1956).

* In vielen Fällen wäre eine experimentelle Aussage darüber, ob eine Phase metallisch leitend ist, für die Bindungsanalyse höchst wertvoll. Leider gibt es kaum systematische Untersuchungen darüber.

7. *Vgl. Strukturbericht* Bd. I, Leipzig (1931).
8. KRONMÜLLER, H. Stuttgarter Diplomarbeit (1956). VGL. A. SEEGER und BROSS H. *Z. Physik* **145,** 161 (1956).
9. VGL. U. DEHLINGER. *Theoretische Metallkunde,* § 14. Berlin-Göttingen-Heidelberg (1955).
10. DEHLINGER U. und SCHULZE G. E. R. *Z. Metallkde* **33,** 157 (1941).
11. FRIEDEL J. Thèse, Paris 1953, *Phil. Mag.* **43,** 153 (1952).
12. DEHLINGER U., *Z. Elektrochem.* **46,** 627 (1940).
13. DEHLINGER und PFLEIDERER H. *Z. Metallkde.* **47,** 229 (1956).

LETTER TO THE EDITORS

Exciton doublet splitting in ionic crystals*

(*Received* 10 *October* 1956)

THE doublet structure of exciton lines in alkali halide crystals was attributed by MOTT and GURNEY[1] to the spin-orbit interaction energy of the $P_{1/2}$ and $P_{3/2}$ core states of the excited negative halogen ions associated with the exciton. The observed splittings in several instances are of the order of magnitude expected from the spectroscopy of free halogen atoms. OVERHAUSER[2] has pointed out another splitting mechanism, the transfer of the excited electron from the anion to the nearest neighbor cations. OVERHAUSER has objected to the MOTT-GURNEY splitting mechanism on the ground that in the L-S coupling approximation the first excited configuration p^5s has $L = 1$ and $S = 0,1$ giving the states $^3P_{0,1,2}$ and $^1P_1$; the only allowed electric dipole transition from the $p^6(^1S)$ ground state is to $^1P_1$, so that a doublet structure is not expected here. This statement assumes that the spin-orbit interaction energy is small and, as OVERHAUSER points out, should apply quite well to excitons in BaO associated with excitation of the O^{--} ion.

It is of some interest to consider the situation in detail for the alkali halides. It is well known to spectroscopists that the spectra of rare gas atoms (isoelectronic with negative halogen ions) show marked doublet structure, and it has been recognized that in these atoms j-j coupling is responsible for the doublet structure. An estimate has been made of the relative intensities that might be expected for the first two lines by taking advantage of the similarity of the halogen ion and the rare gas configuration.

CONDON and SHORTLEY[3] have discussed rare gas spectra. Using their results for the eigenvalues of the spin-orbit and electrostatic interactions it was possible to obtain a very simple expression for the ratio of the squares of the matrix elements of the electric dipole moment operator connecting the two $J = 1$ excited states with the ground state:

$$\frac{|\langle {}^1S_0|\boldsymbol{P}|P_{1/2},\, J=1\rangle|^2}{|\langle {}^1S_0|\boldsymbol{P}|P_{3/2},\, J=1\rangle|^2} = \frac{\epsilon_1+2\epsilon_3-3\epsilon_4}{\epsilon_1+2\epsilon_3-3\epsilon_2}$$

where the ϵ's are the observed energy values[4] of the first four excited states of the particular rare gas configuration in question in order of increasing energy, i.e., $^2P_{3/2}$ $J = 2, J = 1$; $^2P_{1/2}$ $J = 0$, $J = 1$. Calculated values of the ratio of intensities of the doublet lines are given below.

Ne I—13·2	Na II—21·0
A I— 4·0	
Kr I— 1·0	Rb II— 1·1
Xe I— 0·8	

For the krypton configuration it is seen that Kr I and Rb II are not dissimilar, giving some hope that Br^- might have an equal intensity doublet structure.

Since the exciton lines are not well separated it is not possible to tell exactly what their actu intensities are, but in the bromides and iodides where they are resolved there appears to be at least qualitative agreement of the observed[5] intensity ratios and those calculated here.

Professor F. A. JENKINS and Dr. CHARLOTTE E. MOORE have kindly helped with references to the spectroscopic literature.

Department of Physics, RICHARD PETERSEN
University of California,
Berkeley, California.

* Assisted by the Office of Naval Research, the Signal Corps, and the Air Force Office of Scientific Research.

REFERENCES

1. MOTT N. F. and GURNEY R. W. *Electronic Processes in Ionic Crystals* (Oxford University Press, London, 1948); see also FRANCK and SCHEIBE *Z. phys. Chemie* **139,** 22 (1928).
2. OVERHAUSER A. W. *Phys. Rev.* **101,** 1702 (1956).
3. CONDON E. U. and SHORTLEY G. H. *Theory of Atomic Spectra* (Cambridge University Press, Cambridge, 1951).
4. U.S. NATIONAL BUREAU OF STANDARDS *Atomic Energy Levels,* (U.S. Government Printing Office, Washington, D.C., Vol. I, 1949, Vol. II, 1952).
5. PICK H. *Landolt-Börnstein Tables, Vol. I, Part IV,* (Springer-Verlag, Berlin, 1955).

VOL. 1 1956/57

L.
/57